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

DEFENCE R&D DÉFENSE

Isothermal Amplification of Microbial Genomic Samples

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

C. Ruttan and D. Mah
Canada West Biosciences

Defence R&D Canada

Technical Memorandum

DRDC Suffield TM 2010-143

October 2010

Canada

Isothermal Amplification of Microbial Genomic Samples

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

C. Ruttan and D. Mah
Canada West Biosciences

Defence R&D Canada – Suffield

Technical Memorandum
DRDC Suffield TM 2010-143
October 2010

Principal Author

Barry N. Ford
Defence Scientist

Approved by

Stephen Bjarnason
Section Head, Casualty Management Section

Approved for release by

Robin Clewley
Acting Chair, Document Review Panel

Abstract

DNA based microbial detection and identification assays generally have sample input requirements which are determined by the assay system. In some cases, samples may lack sufficient amounts of microbial genetic material for analysis. If the microbial sample being analyzed is not, or cannot be cultured in the lab, then some means of amplifying the genetic material is required. This is particularly important if multiple assays are required from a single sample. Ideally, the amplification method should amplify all sequences with equal frequency (uniform effect), be generic (applicable to all types of microbial targets), robust (applicable to a wide variety of different sample types) and should have no negative influence on downstream analysis. During development of a new microarray design for microbial fingerprinting, it became apparent that for some samples of interest, sufficient quantities of purified DNA test material were not available, thus an amplification methodology was required. Prior to this study, no commercial off the shelf (COTS) genomic amplification kit was explicitly designed for analysis of microbial samples. Consequently, a diverse set of genomic DNA samples from a number of bacterial agents were amplified using a COTS kit designed for use with human DNA with the presumption that this kit might be useful for microbial samples. The amplified samples were analyzed on a genomic fingerprinting microarray assay platform developed by DRDC. Non-amplified data were compared to data from amplified samples. The effective ratio of amplification for samples ranged from a few hundred fold to 300,000 fold, comparable to single-gene polymerase chain reaction (PCR) based methods. All products were analyzed using microarray hybridization, and non-amplified data compared to amplified sample data. Some amplification bias in certain genomic regions was observed, but this should not negatively affect detection/identification analysis from a qualitative perspective. The need for a generic amplification tool for low abundance targets is likely to be an ongoing requirement. The technique described will be employed in ongoing research and development efforts.

Résumé

La quantité d'échantillon requise pour les tests de détection et d'identification de microorganismes fondés sur l'ADN dépend habituellement du système utilisé. Dans certains cas, la quantité de matériel génétique dont on dispose ne suffit pas pour le test voulu. Si les microorganismes de l'échantillon à vérifier ne peuvent être cultivés en laboratoire, il faut recourir à un autre moyen pour en amplifier le matériel génétique. Ceci est d'autant plus important si l'on veut faire plusieurs tests avec un seul échantillon. Idéalement, la méthode d'amplification doit permettre d'amplifier toutes les séquences avec la même fréquence (effet uniforme), elle doit être générique (applicable à tous les types de cibles microbiennes), robuste (applicable à divers types d'échantillons) et elle ne doit pas compromettre les analyses subséquentes qu'on voudrait faire avec le matériel amplifié. Durant la mise au point d'un nouveau système de biopuce pour l'identification de microorganismes, nous nous sommes rendu compte que pour certains échantillons, les quantités d'ADN purifié n'étaient pas suffisantes et qu'il fallait donc amplifier le matériel. Avant cette étude, il n'existait pas de trousse commerciale d'amplification génomique spécialement conçue pour l'analyse des échantillons microbiologiques. Nous avons donc amplifié des échantillons d'ADN génomique de différentes bactéries avec une trousse d'amplification

commerciale destinée à l'amplification de l'ADN humain en supposant qu'elle pouvait être utile dans le cas des échantillons microbiologiques. Les échantillons amplifiés ont été analysés sur une plateforme d'empreintes génomiques mise au point par RDDC. Le matériel non amplifié a été comparé à celui des échantillons amplifiés. Les facteurs d'amplification des échantillons variaient de quelques centaines à 300 000, des valeurs comparables à celles obtenues avec l'amplification par PCR d'un seul gène. Tous les produits obtenus ont été analysés par hybridation sur biopuce, et le matériel non amplifié a été comparé à celui des échantillons amplifiés. Nous avons constaté de légers biais d'amplification dans certaines régions génomiques, mais ceux-ci ne devraient pas compromettre la détection ou l'identification, qui sont des analyses qualitatives. Il semble probable qu'on aura de plus en plus besoin d'un outil générique d'amplification de cibles peu abondantes. La technique que nous décrivons ici servira dans les initiatives de recherche et développement.

Executive summary

Isothermal Amplification of Microbial Genomic Samples:

Barry Ford; Doug Bader; Cindy Ruttan; David Mah; DRDC Suffield TM 2010-143; Defence R&D Canada – Suffield; October 2010.

Introduction: DNA-based microbial detection and identification assays require DNA samples with sufficient amounts of material for analysis. Often however, the sample being tested cannot be cultured in the lab, or the sample contains insufficient DNA for routine assays. During evaluation of a new microarray design for microbial fingerprinting, it became apparent that sufficient quantities of purified DNA test material would not be available for some samples, thus an amplification methodology was required. This is a general problem if one wishes to run multiple assays (*i.e.* assay many gene targets) from a single sample.

Under ideal circumstances, the method used to amplify the sample should have a uniform effect on the samples being analyzed (*i.e.* amplify all parts of the sample DNA equally), be applicable to many samples without modification, and should have no negative influence on downstream analysis. In practice this ideal is partially realized. Prior to this study, no COTS genomic amplification kit was explicitly designed for microarray fingerprinting analysis of microbial samples. Using a COTS kit designed for use with human DNA (REPLI-g) with the presumption that this kit might be useful for microbial samples, multiple DNA samples from a wide variety of bacterial sources were amplified. The amplified samples were analyzed on a genomic fingerprinting microarray developed by DRDC.

Results: Quantitative analysis showed that samples were amplified to a similar total yield, irrespective of the input quantity of genomic material. The effective ratio of amplification for samples ranged from a few hundred fold to 300,000 fold. Every amplification reaction yielded more DNA product than was required for microarray analysis, thereby allowing residual DNA to be archived. Microarray analysis of amplified versus non-amplified samples revealed some amplification bias, leading to increased signals for some sequences. All fragments detected in unamplified samples were also detected in amplified samples. In light of this positive only bias, this bias should not negatively affect detection/identification analysis from a qualitative perspective, particularly with consistent application of methods.

Significance: The REPLI-g kit provides a means to analyze samples with low abundance targets that were previously not amenable to microarray analysis and increases the number and types of tests that can be conducted on a single sample. The REPLI-g kit has additional advantages, including incubation at a single temperature (no thermal cycling), universal application, and a relatively simple protocol.

Future plans: The amplification of low abundance targets is likely to be an ongoing requirement. Despite certain caveats in use, the COTS kit used here, or others with similar capabilities, will be employed in ongoing research and development efforts.

Sommaire

Isothermal Amplification of Microbial Genomic Samples:

Barry Ford; Doug Bader; Cindy Ruttan; David Mah; DRDC Suffield TM 2010-143; R & D pour la défense Canada – Suffield; Octobre 2010.

Introduction ou contexte. La détection et l'identification de microorganismes fondées sur l'ADN requièrent de bonnes quantités d'échantillons. Mais il arrive souvent que l'échantillon à analyser ne puisse pas être cultivé en laboratoire ou encore qu'il ne contienne pas suffisamment d'ADN pour les tests de routine. Pendant notre évaluation d'un nouveau type de biopuce pour l'identification de microorganismes, nous nous sommes rendu compte que pour certains échantillons, les quantités d'ADN purifiées ne seraient pas suffisantes. Il nous fallait donc une méthode d'amplification. Ce genre de problème est courant lorsque l'on veut procéder à plusieurs analyses, c'est-à-dire vérifier plusieurs cibles géniques, avec un même échantillon.

Idéalement, la méthode d'amplification doit avoir un effet uniforme (c'est-à-dire qu'elle doit amplifier également toutes les régions génomiques), elle doit pouvoir s'appliquer telle quelle à de nombreux échantillons (sans modifications) et elle ne doit pas compromettre les analyses subséquentes faites avec le matériel amplifié. Avant notre étude, il n'existait pas de trousse commerciale d'amplification génomique spécialement conçue pour l'identification de microorganismes. Nous avons amplifié des échantillons d'ADN génomique de différentes bactéries avec une trousse d'amplification commerciale destinée à l'amplification de l'ADN humain (REPLI-g) en supposant qu'elle pouvait être utile dans le cas des échantillons microbiologiques. Les échantillons amplifiés ont été analysés sur une plateforme d'empreintes génomiques mise au point par RDDC.

Résultats. L'analyse quantitative a montré que l'amplification des échantillons permettait d'obtenir un nombre de copies similaire, quelle que soit la quantité initiale de matériel génomique. Les facteurs d'amplification des échantillons variaient de quelques centaines à 300 000. Chaque réaction d'amplification a produit plus d'ADN que ce qui était nécessaire pour l'analyse avec la biopuce, permettant ainsi de conserver des restes de matériel. L'analyse avec la biopuce du matériel amplifié par rapport au matériel non amplifié a révélé certains biais liés à l'amplification, lesquels ont entraîné des signaux accrus pour certaines séquences. Tous les fragments détectés dans les échantillons de matériel non amplifié ont aussi été détectés dans le matériel amplifié. Étant donné que ce biais est uniquement positif, il ne devrait pas nuire à la détection ou à l'identification, qui sont des analyses qualitatives, particulièrement lorsque les méthodes sont faites de façon uniforme.

Importance. La trousse REPLI-g permet d'analyser, avec une biopuce, des échantillons dans lesquels les cibles sont peu abondantes, ce qui n'était pas possible auparavant. Elle permet également d'augmenter le nombre et le type d'analyses pouvant être faites avec un même échantillon. La trousse REPLI-g présente aussi d'autres avantages, dont l'incubation à une seule température (pas de cycles thermiques), une application universelle et une marche à suivre relativement simple.

Perspectives. Il semble probable qu'on aura de plus en plus besoin d'amplifier des cibles peu abondantes. Malgré certaines limites dans son utilisation, la trousse commerciale utilisée ici, ou d'autres permettant d'obtenir des résultats semblables, seront employées dans les initiatives de recherche et développement.

This page intentionally left blank.

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 from various strains used in this study.

1 Introduction

Purified genetic material obtained directly from a pure colony is a definitive resource for detection and subsequent identification of microorganisms using genetic techniques. The amount of purified DNA required for analysis is dependent upon the analytical method being used and can range from picogram and even subpicogram quantities for PCR and similar target amplification techniques, to 10-15 µg of pure DNA for microarray-based technologies. During evaluation of a new microarray design for microbial fingerprinting, it became apparent for some samples of interest that sufficient quantities of purified DNA test material (5 µg) were not available, thus an amplification methodology was required. Ideal amplification techniques should amplify all sample templates, and all sequences within a template to a similar degree. In practice such ideal outcomes are unlikely, but a compromise of amplification, utility, and function can be achieved.

A number of commercial-off-the-shelf (COTS) kits were available for general genomic amplification of eukaryotic genomes, but none of these had been developed for prokaryotic fingerprinting, nor were any recommended by the microarray supplier (Affymetrix, Santa Clara, CA) for this purpose. The REPLI-g whole genome amplification kit (QIAGEN, Valencia, CA) was recommended by Affymetrix for single-nucleotide polymorphism microarray analysis of human genomes, and was thus trialed for the purposes of genomic fingerprinting of microbial samples. REPLI-g has not been previously been applied to microbial genomic amplification, so a comparative analysis of amplified versus non-amplified samples on the microarray system was undertaken.

REPLI-g utilizes isothermal amplification based upon the method of Dean et al. [1]. A "random" hexameric oligonucleotide mixture is used to prime DNA synthesis by a highly processive DNA polymerase, in this case DNA polymerase of phage ϕ 29. The polymerase has 3'-5' exonuclease activity to ensure high fidelity. The reaction mixture also contains pyrophosphatase which serves to drive the polymerase reaction forward by degrading released pyrophosphate to free phosphate. The REPLI-g method uses multiple displacement amplification (MDA), which eliminates the requirement for thermal cycling [1, 2] and thus the reaction can be performed at a fixed temperature. This is advantageous in that primer-specific annealing temperatures are not needed, and, at the relatively low incubation temperature, imperfect matches to the primers are also amplifiable, maximizing the general amplification properties of the reaction. As each new strand is synthesized, it displaces a pre-existing strand, which in turn becomes a substrate for priming by the free primer in solution, initiating a new round of synthesis. This process continues until primer is depleted in the reaction to the point where new strands can no longer be efficiently primed. The processive nature of the polymerase means that most initiated strands will continue until the polymerase runs off the end of the template fragment. Thus the product lengths can be very long compared to other amplification methods using chain-terminating thermal denaturation. Typical amplification runs produce a substantial amplification of total DNA, with output to input ratios ranging from 200 to 300,000. Amplified DNA may be used in subsequent labelling and hybridization work. Comparison of amplified to unamplified DNA in microarray experiments showed that genomic DNA amplification detected some target sequences not detected in unamplified samples. The primary advantage of the technique is that very small amounts of sample DNA are required for subsequent analysis.

2 Methods

2.1 DNA amplification

Purified genomic DNA used in this study was obtained from DRDC Suffield stocks, CFIA Lethbridge, and a commercial supplier (Cedarlane®, Burlington, ON, Canada)) (Annex A). Amplification reactions were performed on two independent replicates of each DNA sample, using the REPLI-g kit (Qiagen, Valencia, CA, USA). A schematic description of the REPLI-g amplification reaction method is show in Figure 1. Detailed methods are found an Annex B. The REPLI-g system has the advantage that all the steps are carried out in a single tube. After the amplification reaction was complete, DNA was quantitated by optical density at 260 nm ($OD_{260/280}$; See List of symbols).

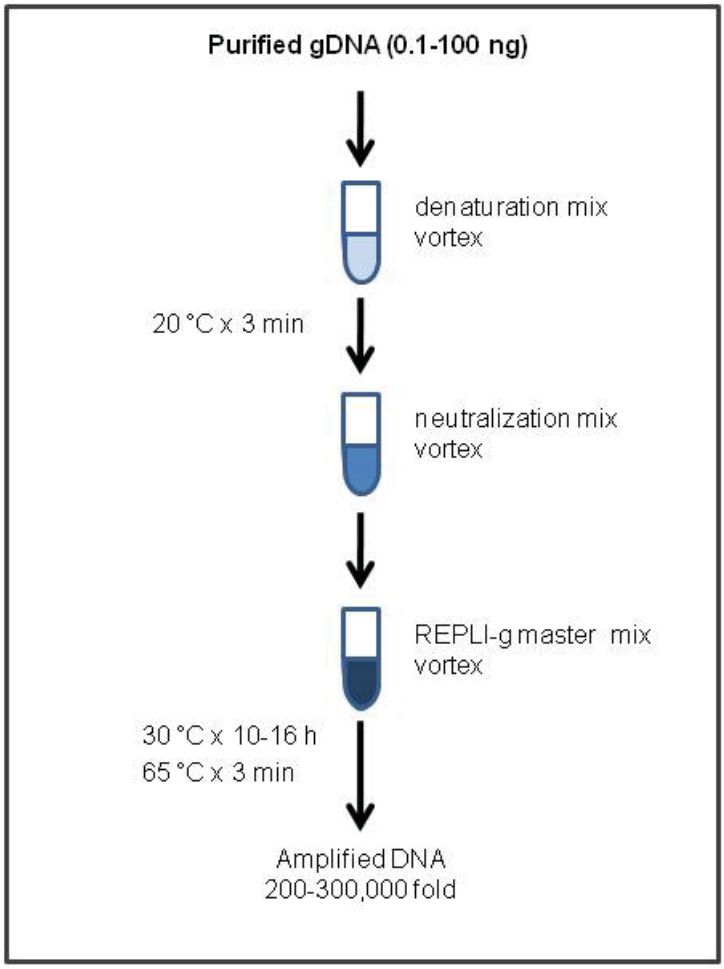

Figure 1: General REPLI-g method

2.2 Microarray hybridization and data analysis

Custom designed microarrays, fabricated by Affymetrix (Santa Clara, CA), were used to evaluate whether hybridization with amplified genomic DNA produced results different from native unamplified DNA. Microarrays were prepared and hybridized according to the manufacturer's recommended method. Microarrays were scanned using the Affymetrix 3000G scanner. Images were digitized automatically, and the resulting data collated into a spreadsheet for analysis. For each unamplified sample, two independent arrays were run. For the amplification samples, single arrays for each amplification reaction for each sample were performed.

During data reduction, locator signals provided on the microarray (used for alignment of the array images) were pruned out of the raw data. Microarray signal data were not scaled for routine analysis. Where scaling was used in a few cases, simple factor scaling to normalize raw data average values was employed. For purposes of discriminating positive from background signals, a conservative arbitrary cut-off of 0.5% (325) of maximal intensity (65535) was used. In principal, the background signal of true negative targets (intensity ranging from 26 to 60 by examination), supplied by the array manufacturer could be used as a minimum cut-off, with greater sensitivity to small signals, but more false differences. The cut-off of 325 was thus approximately tenfold above the typical background signal level. Also removed from the comparison were signal sets (i.e. all signals from a row of aligned data) where the average signal did not exceed the cut-off value.

Qualitative comparison of microarray data was facilitated by conditional formatting of the spreadsheet data, based on signal intensity, followed by manual inspection for differential signals. Differential signals were also detected by statistical analysis, using the Students-t test. Signal sets differing significantly between amplified and unamplified DNA samples were extracted and summarized.

3 Results

Microbial genomic DNA samples were successfully amplified using the REPLI-g system. Input quantities of genomic DNA ranged from 0.1 ng to 113 ng (1000 fold range). Amplified yields ranged from 11,750 ng to 35,875 ng (3 fold range) with an average amplification ratio (yield to input) of 1900 fold. Table 1 contains a subset of the results obtained. Annex A contains a complete listing of the amplification results. Table 1 also indicates the GC content (as percentage) of the microbial genome, to facilitate analysis of whether GC rich genomes might be more difficult to amplify.

Table 1: Input versus amplified DNA yields for selected strains

Sample	% G:C	DNA input (ng)	DNA yield (ng)	Amplification Ratio
S. aureus Z1	32.8	100.2	16000	160
E. coli JM109	50.8	72.5	20625	284
Y. enterocolitica YE-D3	47.3	29.3	18125	618
B. anthracis RP42	35.4	11.3	19250	1704
Y. pestis PX14-3	47.6	11.0	16625	1511
B. mallei ATCC 23344	38.2	1.8	22875	12432
F. tularensis tularensis (A I) Swed1	32.2	0.7	16375	24440
Y. frederiksenii CCRI 14915	~48	0.1	35875	344952

It was apparent within the range of DNA quantities tested that the yield of amplified DNA in terms of total weight (ng), was not a linear function of amount of input DNA, as shown in Figure 2. The amplification ratio (yield divided by input DNA) however was negatively correlated to DNA input quantity. The lower the amount of input DNA, the greater the relative amplification ratio (Figure 3). The declining amplification ratio likely reflects the depletion of the reagents in the reaction mixture. Input amounts as low as 0.1 ng generated enough amplified material to run 6-7 microchips. This reduces the need for large amounts of starting material in order to generate enough material to run on the chip.

The microarray data revealed 221077 target signals remaining after image alignment locator signals were removed. The number of signal intensities on each microarray which exceeded the cut-off minimum value (0.5% of signal maximum) varied depending on the DNA sample analyzed.

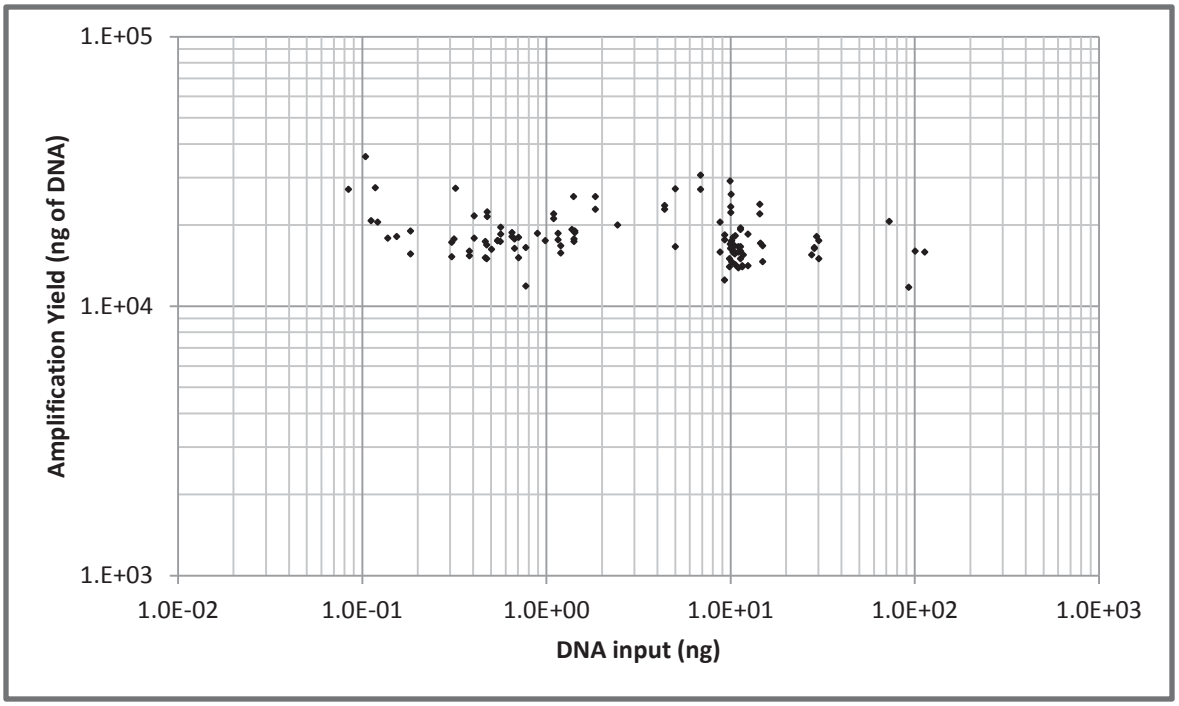

Figure 2: Amplified DNA yield versus input DNA

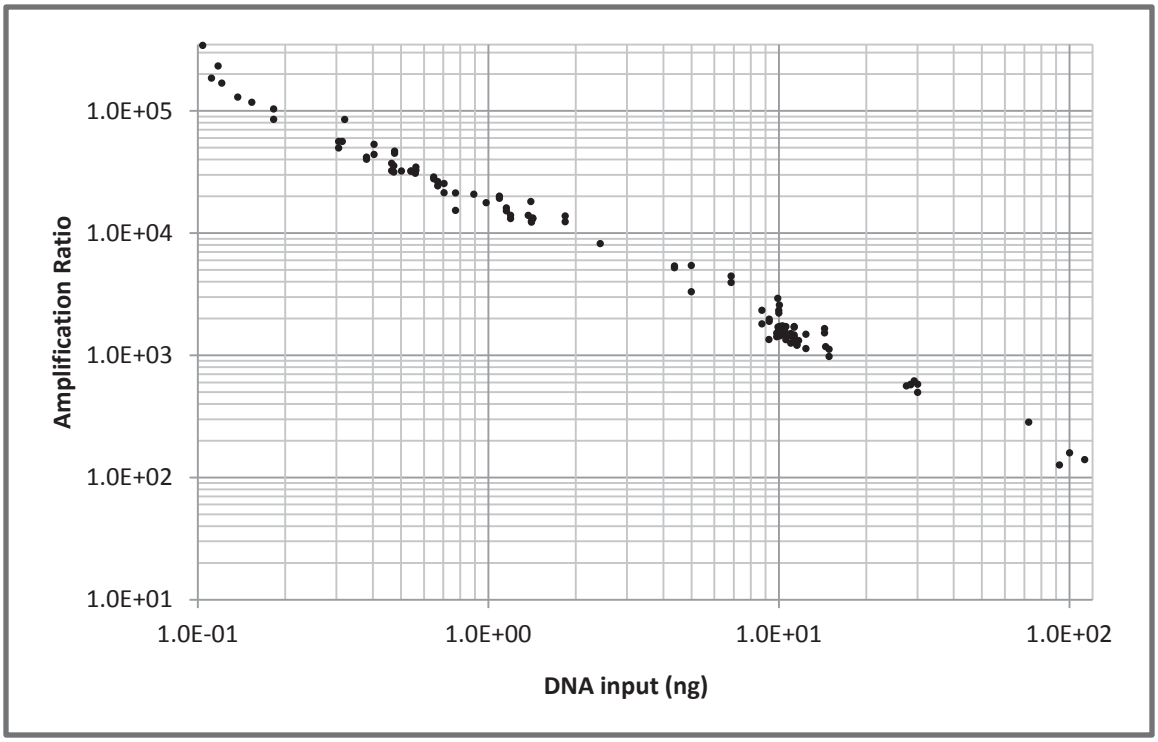

Figure 3: Amplification ratio versus input DNA

For example, in the signal data for *E.coli*, 23817 signals had an average value greater than the cut-off minimum intensity (Table 2). After REPLI-g amplification of *E.coli* genomic DNA, 89 target signals were significantly different in the amplified-DNA versus the unamplified-DNA data. Sixty-nine of these signals represented intensity changes of approximately two-fold or less. Thirty-two accessionable sequence targets were differentially detected in the amplified samples, six of which are full-length *E.coli* sequences (i.e. 25 base pair complete matches) as shown in Table 3. Of these sequences, one was detected without amplification (i.e. the unamplified signal was above the cut-off limit). Eleven differential signals corresponded to targets designed to detect *Shigella flexneri*, which, on further review, also detect *E. coli* in NCBI Entrez-Gene [3]. The best matches to *E. coli* in the remaining differential signals (as detected by blast searching within NCBI Entrez-Gene) comprise, at most, 14 base pairs with complete sequence identity.

Table 2: Signal intensity comparison

Sample	Signal average > minimum	p < 0.05 Amplified vs. Original *	% differential signals	% missed in unamplified
E. coli	23817	89	0.4	0.05
B. anthracis	13167	2663	20.2	1.6
Y. enterocolitica	13363	1387	10.3	26
Y. pestis	8493	933	10.9	65

* Students-t test of two replicates of the native sample, and two independent amplification reactions.

Table 3 : E. coli microarray signals after amplification

Probe ID	Target Sequence	Original Signal	Amplified Signal
12290	TACCGTTCCTGTCACCTTGTCGAAG	143	2395
34038	GCTGCGTCTTGTATGGTTACGAGTG	176	1364
42596	CAGCGCGAACGTGACTCGCAGTTTT	164	2045
60755	AGATTAGCGCGAAAATTATCATCGT	68	2233
62413	CCTTTATAGAAGTAGGCGTCATGGG	446	1958
68696	TCTTCCCGAGTTGAATTGAGGACAT	318	2267

In the *B. anthracis* data, 2663 target spots differed significantly between unamplified and amplified samples. Of 2742 targets specifically designed to detect *B. anthracis*, 172 were significantly different in the amplified versus the unamplified samples, but only 16 were above the cut-off limit (i.e. signal >325) either set. A single target (probe number 80088) yielded a smaller signal in the amplified than in the unamplified sample. This sequence corresponds to an unnamed *B.anthracis* predicted protein, and was found by blast searching at NCBI to be present in *Bacillus thuringiensis* and *Bacillus cereus* as well.

In the case of *Y. enterocolitica*, 363 targets out of 13363 would not have been detected above the minimum cut-off without DNA amplification. This amounts to 26% of the signals found to be significantly different in the amplified sample relative to the unamplified sample. For *Y. pestis*, 607 of 933 differential signals (65%) would not have been detected as positive without amplification.

In all the sample sets analyzed, the amplified DNA yielded more positive signals. In no sample set were signals detected in the unamplified samples, that were not also detected in the amplified samples.

4 Discussion

Genomic DNA amplification is one solution to the problem of insufficient sample DNA for complete analysis by genetic techniques. An ideal amplification method should be species-independent, robust with respect to genomc composition, and irrespective of technology, would be completely unbiased in terms of output relative to input. Thus, the amplification of sample DNA should be uniform, and not lead to changes in the relative representation of sequences in a sample (i.e. all input sequences would be equally amplified). This ideal amplification is unlikely to be achieved, in part due to the diverse composition of DNA sequences within the genome, and the variation in gneome size between organism of interest. Nevertheless, for the purposes of a genomic sample amplification tool for routine applications, a practical compromise of amplification yield, robustness, and limited amplification bias can be accepted.

In the case of the amplified DNA samples shown in Annex A, a wide diversity of microbial genera and species with genomic G:C contents ranging from 32% to 50% were successfully amplified in replicate. A characteristic property of the REPLI-g reaction is that total yield is quite consistent between samples (limited by the reagent mixtures), irrespective of the amount of input DNA. For the genomic samples analyzed in detail in this report, all samples amplified to similar net yields, sufficient for multiple microarray analyses. This similarity in final yield is likely driven to some degree by titration of the reagents in samples with more input DNA (i.e. shorter effective reaction times) and longer effective reaction times in samples with lesser inputs. Thus the isothermal method appears to be quite robust to different genomic compositions, supporting its generic utility. This has the effect of normalizing sample DNA quantity from different reactions, simplifying application of subsequent analytical methods. In addition, it has been reported that isothermal methods are less susceptible to inhibitory contaminants than PCR methods [2], although this was not tested in this work.

For the purposes of whole-genome amplification, non-cycling amplification is designed to amplify the entire sample DNA content with as little differential as possible. The REPLI-g amplification system uses random hexamer primers at non-discriminatory annealing temperature, and thus in principle, should amplify a wide diversity of sequences with roughly similar efficiencies. As the data show however, amplification bias does occur with the isothermal method, at least as implemented by the REPLI-g kit. For the purposes of detection and identification based on multiple targets scored as present or absent, such as a genomic fingerprinting microarray, positive bias in amplification of sample genomic materials should not be overly problematic as long as the method is reproducible.

Microarray hybridization revealed differential amplification of sequences in the input DNA as was apparent in the microarray signal data. For some samples (e.g. *E. coli*) the difference was very small (less than 1% of total signal). In the case of *Y. enterocolitica*, 65% of the differential signals would not have been detected without amplification. In practice, amplification bias at the genome level may be reproducible at the experimental level as users become experienced using this method, such that any sequence tract in a given genome which is over or under-amplified, will be consistently over or under-amplified to a similar degree in replicate reactions. Over-amplification bias of genomic DNA may actually improve the net sensitivity of a detection/identification protocol for some samples, since relatively more detectable signals will be detected for comparison purposes.

Compared to thermal cycling methods (e.g. PCR), isothermal amplification techniques using random primers have the putative advantages of generic application, simpler design, and less bias in amplifying all regions of the genome [2]. PCR methods generally exploit sequence-specific primers under stringent annealing and extension temperatures, which engender high specificity in the amplification reaction. PCR is by design an attempt to amplify specific sequences away from the bulk of the DNA sample. PCR amplification methods in particular are intrinsically positively biased with respect to overall genome content, and this bias is actually exploited to maximize sensitivity of detection analysis. It is notable that no signals detected in the unamplified samples were **not** detected in the amplified data. In other words, the bias in signal detection, and by extension, sample DNA amplification, was uniformly "positve" under these experimental conditions.

References

[1] Dean, F.B., Hosono, S., Fang, L., Wu, X., Faruqi, A.F., Bray-Ward, P., Sun, Z., Zong. Q., Du, Y., Du, J., Dirscoll, M., Song, W., Kingsmore, S.F., Egholm, M., and Lasken, R.S. (2002), Comprehensive human genome amplification using multiple displacement amplification, Proc. Natl. Acad. Sci., USA, 99, 5261.

[2] Hosono, S., Faruqi, A.F., Dean, F.B., Du, Y., Sun, Z., Wu, X., Du, J., Kingsmore, S.F., Engholm, M., and Lasken, R.S. (2003), Unbiased whole-genome amplification directly from clinical samples, Genome Res., 13, 964.

[3] Anon. (n.d.), Entrez-Gene (online), National Center for Biotechnology Information, U.S. National Library of Medicine, http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene (Access date 10 Jan. 2010).

Annex A Supplementary Data

Genomic DNA Amplification

The following table includes all amplified samples, including replicates.

Microbe	Strain or designation	Source	DNA input ng	Yield ng	Amp Ratio
Acinetobacter baumannii	ATCC 17978	ATCC	11	15875	1443
Acinetobacter baumannii	ATCC 17978	ATCC	11	13875	1261
Bacillus anthracis	RP42	DRDC	11.3	19250	1704
Bacillus anthracis	RP42	DRDC	11.3	19500	1726
Bacillus anthracis	NH	DRDC	0.5	17500	32288
Bacillus anthracis	Vollum	DRDC	0.6	18750	28891
Bacillus anthracis	Ames PLG6	DRDC	6.9	27125	3957
Bacillus anthracis	ACB	DRDC	0.6	19625	34827
Bacillus anthracis	NH	DRDC	0.5	17500	32288
Bacillus anthracis	Vollum	DRDC	0.6	18125	27928
Bacillus anthracis	Ames PLG6	DRDC	6.9	30625	4468
Bacillus anthracis	ACB	DRDC	0.6	18500	32831
Bacillus anthracis	94188c	DRDC	0.3	17750	56439
Bacillus cereus	ATCC 11778	CFIA	9.8	14000	1423
Bacillus cereus	ATCC 11778	CFIA	9.8	15000	1524
Bacillus cereus	ATCC 10987 (CR)	Cedarlane	9.9	17000	1710
Bartonella henselae	ATCC 49882	Cedarlane	8.8	20500	2343
Bartonella henselae	ATCC 49882	Cedarlane	8.8	15875	1814
Bordetella pertussis	ATCC BAA-589	Cedarlane	30.1	15000	499
Bordetella pertussis	ATCC BAA-589	Cedarlane	30.1	17500	582
Burkholderia pseudomallei	Env-81 "7"	DRDC	0.4	21625	53461
Burkholderia pseudomallei	Env-FB20 "5"	DRDC	0.5	22375	47006
Burkholderia mallei	ATCC 1053 "8"	DRDC	1.8	22875	12432
Burkholderia mallei	ATCC 23344 "10"	DRDC	1.1	22000	20137
Burkholderia mallei	ATCC 1053 "8"	DRDC	1.8	25500	13859
Burkholderia mallei	ATCC 23344 "10"	DRDC	1.1	21125	19336
Burkholderia pseudomallei	Env-81 "7"	DRDC	0.4	17875	44190
Burkholderia pseudomallei	Env-FB20 "5"	DRDC	0.5	21500	45168
Campylobacter jejuni	ATCC 700819	Cedarlane	28.4	16375	576
Campylobacter jejuni	ATCC 700819	Cedarlane	28.4	16500	580

Clostridium perfringens	ATCC 13124	Cedarlane	5	27250	5450
Clostridium perfringens	ATCC 13124	Cedarlane	5	16625	3325
Enterococcus faecalis	ATCC 29212	CFIA	10.3	15875	1541
Enterococcus faecalis	ATCC 29212	CFIA	10.3	18000	1748
Escherichia coli	JM109	DRDC	72.5	20625	284
Escherichia coli	JM109	DRDC	14.5	17125	1181
Escherichia coli	0517:H7 EDL933	CFIA	10.1	17375	1724
Escherichia coli	ATCC 25922	CFIA	11.3	16625	1471
Escherichia coli	0517:H7 EDL933	CFIA	10.1	16375	1625
Escherichia coli	ATCC 25922	CFIA	11.3	15000	1327
Francisella tularensis holarctica	Swed 4Q	DRDC	1.4	19250	14020
Francisella tularensis holarctica	Swed 4Q	DRDC	0.1	17875	130189
Francisella tularensis holarctica	Swed 6Q	DRDC	1.4	17750	12602
Francisella tularensis holarctica	Swed 6Q	DRDC	1.4	17375	12336
Francisella tularensis holarctica	Swed 3	DRDC	0.5	16250	32338
Francisella tularensis holarctica	Swed 9	DRDC	2.4	20000	8234
Francisella tularensis holarctica	Swed 10	DRDC	1.4	25500	18195
Francisella tularensis holarctica	FT-65-4	DRDC	1	17500	17794
Francisella tularensis holarctica	FT-67-4	DRDC	0.9	18625	20880
Francisella tularensis mediasiatica	Swed 8-6	DRDC	1.2	17625	15266
Francisella tularensis mediasiatica	Swed 8-6	DRDC	1.2	18625	16133
*Francisella tularensis tularensis (*A I)	Swed 1	DRDC	0.7	16375	24440
*Francisella tularensis tularensis (*A I*)*	Swed 2Q	DRDC	0.5	15000	31712
Francisella tularensis tularensis (A I)	Swed 1	DRDC	0.7	17750	26493
Francisella tularensis tularensis (A I)	Swed 2Q	DRDC	0.5	16875	35677
Francisella tularensis tularensis (A I)	Swed 2 (IDI)	DRDC	0.6	17375	30971
Francisella tularensis tularensis (A II)	Swed 7Q	DRDC	1.4	19000	13347
Francisella tularensis tularensis (A II)	Swed 7Q	DRDC	1.4	18750	13172
Haemophilus influenzae	ATCC 51907	Cedarlane	10	16375	1638
Haemophilus influenzae	ATCC 51907	Cedarlane	10	16375	1638
Listeria monocytogenes	NTCC 7933	CFIA	12.4	14125	1139
Listeria monocytogenes	ATCC 15313	CFIA	10.5	15875	1515
Listeria monocytogenes	NTCC 7933	CFIA	12.4	18500	1492
Listeria monocytogenes	ATCC 15313	CFIA	10.5	16750	1598
Mycobacterium BCG	ATCC 19015	Cedarlane	4.4	23625	5400
Mycobacterium BCG	ATCC 19015	Cedarlane	4.4	22875	5229
Mycoplasma pneumoniae FH	ATCC 15531	Cedarlane	14.4	22000	1530

Mycoplasma pneumoniae FH	ATCC 15531	Cedarlane	14.4	23875	1661
Mycoplasma pneumoniae FH	ATCC 15531	Cedarlane	10	23375	2338
Pseudomonas aeruginosa	ATCC 27853	CFIA	113	15875	140
Pseudomonas aeruginosa	ATCC 27853	CFIA	11.3	16000	1416
Salmonella typhimurium	71-471	CFIA	10.1	16500	1633
Salmonella typhimurium	71-471	CFIA	10.1	14625	1448
Shigella dysenteriae	ATCC 11835	CFIA	11.6	14000	1212
Shigella dysenteriae	ATCC 11835	CFIA	11.6	14125	1223
Staphylococcus aureus	Z1	CFIA	100.2	16000	160
Staphylococcus aureus	Z1	CFIA	10.5	15625	1486
Streptococcus pyogenes	ATCC 19615	CFIA	9.3	18375	1982
Streptococcus pyogenes	ATCC 19615	CFIA	9.3	17625	1901
Vibrio vulnificus	Z86	CFIA	92.5	11750	127
Vibrio vulnificus	Z86	CFIA	9.3	12500	1351
Yersinia aldovae	ATCC 35237	DRDC	0.7	15125	21469
Yersinia aldovae	ATCC 35237	DRDC	0.7	18000	25550
Yersinia bercovieri	CCRI 14920	DRDC	0.2	18125	118078
Yersinia enterocolitica	YE-D3	DRDC	29.3	18125	620
Yersinia enterocolitica	YE-D3	DRDC	11.7	15500	1325
Yersinia enterocolitica	CCUG 33553	DRDC	0.3	17250	56465
Yersinia enterocolitica	#14 field strain	CFIA	14.9	16750	1124
Yersinia enterocolitica	CCUG 33553	DRDC	0.3	15250	49918
Yersinia enterocolitica	#14 field strain	CFIA	14.9	14625	982
Yersinia enterocolitica	CCUG 31436	DRDC	0.1	27500	234043
Yersinia enterocolitica	#7 field strain	CFIA	10	17000	1701
Yersinia enterocolitica	ATCC 23715	CFIA	10	22250	2224
Yersinia enterocolitica	gDNA	DRDC	10.1	26000	2587
Yersinia frederiksenii	CCRI 14915	DRDC	0.1	35875	344952
Yersinia intermedia	ATCC 33648	DRDC	0.1	27125	322917
Yersinia kristensenii	ATCC 33638	DRDC	0.1	20500	169421
Yersinia mollaretii	ATCC 43969	DRDC	0.1	20750	186099
Yersinia pestis	PX14-3	DRDC	27.5	15500	564
Yersinia pestis	PX14-3	DRDC	11	16625	1511
Yersinia pestis	CO92	DRDC	0.8	11875	15392
Yersinia pestis	C12	DRDC	1.2	15750	13196
Yersinia pestis	GB	DRDC	0.2	15625	85616
Yersinia pestis	CO92	DRDC	0.8	16500	21387
Yersinia pestis	C12	DRDC	1.2	16750	14034
Yersinia pestis	GB	DRDC	0.2	19000	104110
Yersinia pestis	PP65-BC YC-1D (BF)	DRDC	9.9	29125	2936
Yersinia pseudotuberculosis	ATCC 29833	CFIA	10.6	14250	1347

Yersinia pseudotuberculosis	ATCC 6902	DRDC	0.5	15125	32527
Yersinia pseudotuberculosis	ATCC 29833	CFIA	10.6	18250	1725
Yersinia pseudotuberculosis	ATCC 6902	DRDC	0.5	17375	37366
Yersinia pseudotuberculosis	ATCC 13979 (CR)	DRDC	10.1	17500	1730
Yersinia rohdei	CCRI 14919	DRDC	0.3	27375	85413
Yersinia ruckerii	ATCC 29473	DRDC	0.4	15375	40354

Annex B REPLI-g Detailed Method

This protocol is optimized for whole genome amplification from >10 ng of purified genomic DNA template. The template DNA should be suspended in TE buffer (10 mM Tris, 1 mM EDTA, pH 7.5). Smaller amounts (1–10 ng) of starting material can be used if the DNA is of sufficient quality. For best results, the template DNA should be >2 kb in length with some fragments >10 kb.

REPLI-g Mini DNA Polymerase should be thawed on ice. All other components can be thawed at room temperature (15–25 °C). Buffer D1 and Buffer N1 should not be stored longer than 3 months.

Prepare Buffer DLB by adding 500 µl nuclease-free water to the tube. Mix thoroughly and centrifuge briefly. Note: Reconstituted Buffer DLB can be stored for 6 months at –20 °C. Buffer DLB is pH-labile. Avoid neutralization with CO_2. All buffers and reagents should be vortexed before use to ensure thorough mixing. Set a water bath or heating block to 30 ºC.

Preparations

1. Prepare sufficient Buffer D1 (denaturation buffer) and Buffer N1 (neutralization buffer) for the total number of whole genome amplification reactions. Buffer D1 and Buffer N1 should not be stored longer than 3 months.

2. Place 5 µl template DNA into a microcentrifuge. The amount of template DNA should be >10 ng. A DNA control reaction can be set up using 10 ng (1 µl) control genomic DNA. Adjust the volume with TE to the starting volume of your sample.

3. Add 5 µl Buffer D1 to the DNA. Mix by vortexing and centrifuge briefly.

4. Incubate the samples at room temperature for 3 min.

5. Add 10 µl Buffer N1 to the samples. Mix by vortexing and centrifuge briefly.

6. Thaw REPLI-g Mini DNA Polymerase on ice. Thaw all other components at room temperature, vortex, then centrifuge briefly. The REPLI-g Mini Reaction Buffer may form a precipitate after thawing. The precipitate will dissolve by vortexing for 10 s.

7. Prepare a master mix on ice. Mix and centrifuge briefly. Important: Add the master mix components in the order listed. After the addition of water and REPLI-g Mini Reaction Buffer, briefly vortex and centrifuge the mixture before the addition of REPLI-g Mini DNA Polymerase. The master mix should be kept on ice and used immediately upon addition of the REPLI-g Mini DNA Polymerase.

Preparation of Buffer D1

Volumes given are suitable for up to 15 reactions.

Reconstituted Buffer DLB	9 µl
Nuclease-free water	32 µl
Total volume	41 µl

Preparation of Buffer N1

Volumes given are suitable for up to 15 reactions.

Stop solution	12 µl
Nuclease-free water	68 µl
Total volume	80 µl

Protocol

8. Add 40 µl of the master mix to 20 µl of denatured DNA (step 5).

9. Incubate at 30 °C for 10–16 h. Maximum DNA yield is achieved using an incubation time of 16 h. After incubation at 30 °C, heat the water bath or heating block up to 65 °C if the same water bath or heating block will be used in step 10. Note: If a thermal cycler is used with a heated lid, temperature of the lid should be set to 70 °C.

10. Inactivate REPLI-g Mini DNA Polymerase by heating the sample for 3 min at 65 °C.

11. If performing PCR analysis, dilute the amplified DNA 1:20 and use 3 µl of diluted DNA for each PCR. Note: For dilution, add 2 µl amplified DNA to 38 µl water or TE. Use 3 µl of the diluted DNA for each PCR.

12. Store amplified DNA at 4 °C for short-term storage or –20 °C for long-term storage. DNA amplified using the REPLI-g kit should be treated as genomic DNA with minimal freeze-thaw cycles. Storage of nucleic acids at low concentration over a long period of time may result in acid hydrolysis. Storage of nucleic acids at a concentration of at least 100 ng/µl is recommended.

List of symbols/abbreviations/acronyms/initialisms

APE	apurinic endonuclease; cleaves DNA adjacent to apurininc sites
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
CFIA	Canadian Food Inspection Agency
COTS	commercial off the shelf
DNA	deoxyribonucleic acid
DRDC	Defence Research & Development Canada
HPT	hypoxanthine guanine phosphoribosyltransferase
mM	millimolar
MDA	multiple displacement amplification
NCBI	National Center for Biotechnology Information (also referred to as Genbank)
$OD_{260/280}$	quantity and relative purity of DNA samples can be assessed by measuring the sample optical density at 260 and 280 nm, then generating the ratio OD260:OD280. 1 OD_{260} corresponds to 50 µg of DNA, whereas $OD_{260/280}$ ratios in the range of 1.5-1.8 indicate good DNA sample purity.
PCR	polymerase chain reaction, a standard DNA amplification technique
PERL	a high level programming language for scanning text files, extracting data, and generating reports from the data
R&D	Research & Development
SNP	single nucleotide polymorphism; a sequence variant at one base position which may be different between populations or individuals
TDT	terminal deoxynucleotidyl transferase

This page intentionally left blank.

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 REVIEW: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.)

Isothermal Amplification of Microbial Genomic Samples:

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

Ford, B.N.; Bader, D.; Ruttan, C.;Mah, D.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)	6b. NO. OF REFS (Total cited in document.)
October 2010	29	3

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 Memorandum

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.)	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)
DRDC Suffield TM 2010-143	

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.)

DNA based microbial detection and identification assays generally have sample input requirements which are determined by the assay system. In some cases, samples may lack sufficient amounts of microbial genetic material for analysis. If the microbial sample being analyzed is not, or cannot be cultured in the lab, then some means of amplifying the genetic material is required. This is particularly important if multiple assays are required from a single sample. Ideally, the amplification method should amplify all sequences with equal frequency (uniform effect), be generic (applicable to all types of microbial targets), be robust (applicable to a wide variety of different sample types) and should have no negative influence on downstream analysis. During development of a new microarray design for microbial fingerprinting, it became apparent that for some samples of interest, sufficient quantities of purified DNA test material were not available, thus an amplification methodology was required. Prior to this study, no commercial off the shelf (COTS) genomic amplification kit was explicitly designed for analysis of microbial samples. Consequently, a diverse set of genomic DNA samples from a number of bacterial agents were amplified using a COTS kit designed for use with human DNA with the presumption that this kit might be useful for microbial samples. The amplified samples were analyzed on a genomic fingerprinting microarray assay platform developed by DRDC. Non-amplified data were compared to data from amplified samples. The effective ratio of amplification for samples ranged from a few hundred fold to 300,000 fold, comparable to single-gene polymerase chain reaction (PCR) based methods. All products were analyzed using microarray hybridization, and non-amplified data compared to amplified sample data. Some amplification bias in certain genomic regions was observed, but this should not negatively affect detection/identification analysis from a qualitative perspective. The need for a generic amplification tool for low abundance targets is likely to be an ongoing requirement. The technique described will be employed in ongoing research and development efforts.

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.)

microbial genotyping, genomic fingerprinting, detection, identification, microarray, isothermal amplification

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