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

DRDC | RDDC
technologysciencetechnologie

Genetic confirmation of DRDC – Suffield Research Centre biological field simulants for the New Substances Notification Program

Doug Bader
Glen Fisher
Susan Rowsell
DRDC – Suffield Research Centre

Defence Research and Development Canada
Scientific Report
DRDC-RDDC-2018-R144
June 2018

Canada

IMPORTANT INFORMATIVE STATEMENTS

This document was reviewed for Controlled Goods by Defence Research and Development Canada (DRDC) using the Schedule to the *Defence Production Act*.

Endorsement statement: This publication has been peer-reviewed and published by the Editorial Office of Defence Research and Development Canada, an agency of the Department of National Defence of Canada. Inquiries can be sent to: Publications.DRDC-RDDC@drdc-rddc.gc.ca.

Abstract

In order to continue to use non-pathogenic organisms as biological warfare agent simulants for open air release, DRDC – Suffield Research Centre was required by Environment Canada to confirm the identities of the simulants employed and submit detailed information packages on each of the simulants in order to comply with the *New Substances Notification Regulations*. The simulant organisms were assessed by Polymerase Chain Reaction (PCR) and in silico analysis. The simulant organisms included *Pantoea agglomerans* (a Gram-negative bacterial simulant formerly named *Erwinia herbicola*), *Bacillus atrophaeus* (a Gram-positive bacterial spore simulant formerly named *Bacillus globigii*) and Male Specific Coliphage 2 (MS2, a viral simulant) including its *Escherichia coli* host. Published polymerase chain reaction (PCR) assays for each of the organisms were reviewed, selected, modified as required, and put into practice. *In silico* analysis was conducted to address Environment Canada's concern regarding the specificity of the *aro*Q assay to detect the *P. agglomerans* simulant strain and not related *Pantoea* strains, including plant pathogens. To address residual prevalence of the simulants prior to new releases, attempts were made to isolate them from DRDC – Suffield Research Centre field study sites where they were previously used. The PCR identification method for *B. atrophaeus* was used to confirm the identity of spore isolates found in environmental samples. The PCR identification methods for the other simulants were not employed as they were not found in these samples. The data and information presented in this report led to the approval of outdoor aerosol release of MS2/*E.coli* and *P. agglomerans* at DRDC – Suffield Research Centre by Environment Canada. *B. atrophaeus* (BG) was also approved, but was not required to go through the NSN submission process, as it was instead placed on the NSN Program Domestic Substances List due to its historical use by DRDC – Suffield Research Centre.

Significance to defence and security

As a result of this work, Environment Canada has approved outdoor aerosol release of live biological agent simulants at DRDC – Suffield Research Centre which will be used for training and for assessment of detection and protection equipment and protocols.

Résumé

Environnement Canada a demandé au Centre de recherche de Suffield de RDDC de confirmer l'identité des organismes non pathogènes simulant des agents de guerre biologique (« simulants ») qui sont dispersés dans l'atmosphère et de lui envoyer un ensemble de renseignements détaillés sur chacun d'eux afin que le Centre se conforme au *Règlement sur les renseignements concernant les substances nouvelles* et puisse continuer cette pratique. Les simulants ont été évalués par réaction en chaîne de la polymérase (PCR) ou analyse *in silico*. Les organismes simulant trouvés étaient notamment *Pantoea agglomerans* (une bactérie à Gram négatif, auparavant appelée *Erwinia herbicola*), *Bacillus atrophaeus* (une spore bactérienne à Gram positif, auparavant appelée *Bacillus globigii*) et le coliphage mâle spécifique 2 (MS2, un simulant de virus) incluant son hôte *Escherichia coli*. Les essais de réaction en chaîne de la polymérase (PCR) de chacun des organismes ont été étudiés, choisis, modifiés au besoin et mis en pratique. L'analyse *in silico* a été réalisée pour répondre aux préoccupations d'Environnement Canada relatives à la spécificité du test aroQ pour la détection de la souche du simulant *P. agglomerans* plutôt que les souches de *Pantoea* non parentes, y compris les pathogènes végétaux. Pour étudier la question de la prévalence résiduelle des simulants avant une nouvelle dispersion, nous avons tenté de les isoler dans les sites d'étude sur le terrain du centre de recherche de Suffield de RDDC où ils avaient été dispersés antérieurement. La méthode d'identification de *B. atrophaeus* a été utilisée pour confirmer l'identité des isolats de spores trouvés dans les échantillons environnementaux. Nous n'avons pas employé les méthodes d'identification par PCR pour les autres simulants, puisque ceux-ci n'ont pas été retrouvés dans ces échantillons. Les données et les informations présentées dans ce rapport ont mené à l'approbation par Environnement Canada de la dispersion à l'extérieur d'aérosol de MS2/*E. coli* et *P. agglomerans* au Centre de recherche de Suffield de RDDC. Le rejet *B. atrophaeus* (BG) a également été approuvé, mais le processus de divulgation de renseignements concernant les substances nouvelles n'a pas été nécessaire, car cet organisme a plutôt été ajouté à la *Liste intérieure des substances* du Programme des substances nouvelles, à cause de son utilisation de longue date par le Centre de recherche de Suffield de RDDC.

Importance pour la défense et la sécurité

Grâce à ce travail, Environnement Canada a autorisé la dispersion à l'extérieur par le Centre de recherche de Suffield de RDDC de simulants vivants d'agent de guerre biologique qui seront utilisés pour l'entraînement ainsi que l'évaluation du matériel de détection et de l'équipement de protection, et des protocoles.

Table of contents

List of figures

List of tables

Acknowledgements

The authors wish to acknowledge the following people for their contributions: Dave Mader for his technical assistance in performing some of the PCR assay testing; Boyd Garrecht and Laurie Stadnyk for preparing the soil samples and some of the control strains.

This page intentionally left blank.

1 Introduction

DRDC – Suffield Research Centre uses biological agent simulant organisms for training the Canadian Armed Forces and for assessment of equipment and protocols in the areas of detection (e.g., detectors, collectors) and protection (e.g., masks, clothing). While some of this work is conducted in the laboratory, some is conducted in the field, in particular, the DRDC – Suffield Research Centre Experimental Proving Ground (EPG). The simulants that were used in open air releases prior to 2010 included *Pantoea agglomerans*, *Bacillus atrophaeus*, and Male Specific Coliphage 2 with its *Escherichia coli* host. In order to continue to use these biological agent simulants for field studies on the EPG beyond 2010, DRDC – Suffield Research Centre was required to submit a New Substances Notification (NSN) to Environment Canada to meet new environmental regulations. The Centre was requested to use scientific methods to provide objectivity and rigor in addressing the following scientific questions. Is there evidence for the presence of pre-existing simulant biological agents in the EPG test environment; and can the Centre establish analytical methods to qualify the three simulant organisms to be used in future releases?

Pantoea agglomerans is a Gram-negative, aerobic bacillus in the *Enterobacteriaceae* family [1]. It is used at DRDC – Suffield Research Centre as a simulant for Gram-negative human pathogens such as *Brucella* spp., *Burkholderia* spp., *Francisella tularensis*, and *Yersinia pestis*. In 1989, the genus *Pantoea* was established to include several species belonging to the *Erwinia-Enterobacter agglomerans* and *Erwinia milletiae* complex [2]. Currently, the genus *Pantoea* includes seven species and two subspecies, the majority of which are either associated with plants or are plant pathogens [2]. *P. agglomerans* is found in many diverse natural and agricultural habitats and has been isolated from plants, water, animals, and humans [3] and from the feces of humans and animals [4]. *P. agglomerans* has been transformed into two related gall-forming pathovars (*P. agglomerans* pv. gypsophilae and *P. agglomerans* pv. betae) through the acquisition of pPATH plasmids containing a pathogenicity island [3]. However, other *P. agglomerans* strains have biocontrol properties for a variety of bacterial and fungal diseases [5]. For example, *P. agglomerans* strain E325 is the active ingredient in Bloomtime Biological™ FD Biopesticide, used to control fire blight of apple and pear most commonly initiated by *Erwinia amylovora* [6]. Commercial registration of *P. agglomerans* biocontrol products has been hampered as it is considered a Risk Group 2 pathogen due to clinical reports of it acting as an opportunistic pathogen [5]. In terms of human infections, *P. agglomerans* is less often implicated than *Enterobacter aerogenes* or *Enterobacter cloacae* [4]. Two cases of septic monoarthritis have been reported in which *P. agglomerans* was indicated following plant thorn injury (13 year old boy) and a wood sliver injury (36 year old woman with systemic lupus erythematosus) [4]. *P. agglomerans* has been cultured from patients at various locations including sterile sites (blood stream, abscesses, joints/bones, urinary tract, peritoneum, and thorax), sputum, urine, and oropharyngeal [1]. However, genotypic analysis of plant and clinical strains from various culture collections found many clinical strains to be improperly designated as *P. agglomerans*, resulting in taxonomic mischaracterization [5]. In many mischaracterized cases, Koch's postulates are often not fulfilled and the clinical strains are rarely retained for genetic confirmation; these concerns, combined with the polymicrobial nature of *P. agglomerans*, need to be taken into consideration before preventing legitimate use of beneficial *P. agglomerans* as biocontrols [5]. Other subject matter experts, however, have suggested that all *P. agglomerans* strains might possess indistinguishable virulence potential [7], which would therefore argue for caution in using any *P. agglomerans* strain as a biocontrol agent. The *P. agglomerans* strain used by DRDC – Suffield Research Centre and submitted to Environment Canada for the NSN submission is *P. agglomerans* ATCC 33243, formerly referred to as *Erwinia herbicola* [8].

Bacillus atrophaeus is a Gram-positive, aerobic, endospore-forming, rod-shaped bacterium that is virtually identical to *Bacillus subtilis* except for the production of pigment on media which contains organic nitrogen [9]. It is found in soil and is non-pathogenic [10]. It has been used for over 60 years in the biodefense community as a simulant for *Bacillus anthracis* and other spore-forming organisms [10]. It has the Tripartite agent designation BG [10], derived from *Bacillus globigii*, the original name given to this organism by Migula in 1900 [11]. It was originally isolated as a variant of *B. subtilis*, distinguishable by the formation of black-tinted pigment on nutrient agar, hence the designation *Bacillus subtilis* var *niger* [10].[1] Two strains from the Bacon Laboratories, the "red" strain and the "brown" strain, were allocated to *B. subtilis* var *niger* and designated as NRS-1221A and NRS-1221B, respectively [11]. In 1989, the black-pigment-producing strains were evaluated on the basis of pigment production on two different media and DNA-DNA hybridization studies by Nakamura [11]. From this analysis, three groups emerged. One group produced no pigment on either media and included the type strain *B. subtilis*. Another group was a pigment-forming variant, but belonged to *B. subtilis* based on the high DNA–DNA similarities. The third group produced a brownish-black pigment on one medium and a brown pigment on the other, but showed low levels of DNA–DNA homology to the other two groups. This group was designated as a new species, namely, *Bacillus atrophaeus*. In the 1989 study, 21 of 25 strains in this new *B. atrophaeus* group had formerly been designated as *B. subtilis* var *niger*. For many decades, DRDC – Suffield Research Centre has employed "BG powder," freeze-dried *Bacillus subtilis* var *niger* spp. *globigii* (BG) spores prepared by Dugway Proving Grounds (DPG) and obtained in two batches, one apparently prepared in the 1960s and the other in the 1990s. Both batches tend to produce orange-colored colonies on Luria-Bertani (LB) media.

MS2 bacteriophage is a positive-sense, single-stranded RNA (ssRNA) virus with an icosahedral, tail-less capsid [12], whose host organism is *Escherichia coli* [13]. It is a non-enveloped virus with an average diameter of about 28 nm, making it one of the smallest viruses [13]. It binds to the side of bacterial F-pili and lyses the bacterial host after infection [14]. It has been used in place of pathogenic viruses in a wide range of studies, including testing compounds for disinfecting surfaces, studying environmental transport and environmental fate, and as a pathogen simulant in R&D and T&E analysis of methods, systems, and devices for the detection of pathogens in both battlefield and domestic defence scenarios [15].

Escherichia coli K-12S A/λ (F+) is the host strain used to propagate MS2 bacteriophage at DRDC – Suffield Research Centre. This K-12-derived bacterium is Gram-negative, facultative, rod-shaped, and carries the F+ and λ-resistant characters [16]. *E. coli* K-12 and its derivatives are no longer considered hardy in the natural environment due to their extensive use in the laboratory for over 70 years [17]. It is assumed that over the long term, K-12 survival in soil would be very low [17]. These strains, unlike other strains of *E. coli* that are known to be pathogenic to humans, are non-pathogenic. The U.S. Food and Drug Administration states that some non-pathogenic strains can be opportunistic pathogens causing infections in immunocompromised hosts [18]; however, no case of disease has ever been reported through extensive use of K-12 derived *E. coli* strains [19].

The scientific approach that was established and accepted by Environment Canada to address their regulatory requirements involved (i) collecting representative surface soil samples in the EPG from areas that were reasonably expected to harbor simulants from previous releases; (ii) performing sample preparation techniques that promoted the selection and enrichment of each simulant based on differences in their physical, chemical, and/or biological properties; and (iii) qualifying viable simulant populations using simulant-specific Polymerase Chain Reaction (PCR) assays. Although whole genome sequencing

[1] Pigmentation was shown to be susceptible to the culture conditions [11].

(WGS) could have been used to qualify simulant agents rather than PCR, WGS capabilities at the time of this study were not readily available, were very costly, and they would have taken too long to meet the time-sensitive requirement to obtain regulatory approval in time for planned field studies.

PCR assays were used on cultured isolates from soils rather than directly on soils for several reasons. First, the intent of the study was to look for the presence of surviving simulant populations as they would have the potential to have a greater impact on the environment than non-viable populations. Hence culture was used to enrich for viable populations since PCR, by itself, cannot address viability concerns. Second, PCR is usually more effective when conducted on pure cultures than when used directly on soil samples since soils often contain PCR inhibitory components that can reduce the chances of detecting the simulant agents in soil. Third, PCR must be conducted on genetic material contained within the bacterium. In the case of *Bacillus atrophaeus*, it tends to form spores in order to survive in the environment. As such, this simulant in the spore state is more difficult to lyse in order to access the genetic material for PCR. Growing isolated spores on agar plates forces the spores into the vegetative state which improves cell lysis and thus increases the number of starting copies of target sequence for PCR. Finally, agent cultivation before performing PCR allows one to increase the number of copies of target nucleic acid through cellular replication.

SYBR Green dye-uptake polymerase chain reaction (PCR) assays for *Pantoea agglomerans*, *Bacillus atrophaeus*, and MS2 previously developed for the LightCycler® under The Technical Cooperation Program (TTCP) Action Group 43 (Gene Probes) [20] were initially considered for this project but were dismissed based on the fact that they were developed as a quadrilateral effort and may have been subject to possible limitations if data were to be published within the public domain. Consequently, it was decided to utilize real-time, probe-based PCR assays, published in the open literature, for the three DRDC – Suffield Research Centre simulants selected for NSN certification. The added advantage to using probe-based PCR assays allows one to monitor the PCR reaction in real-time and they tend to be more rigorous than SYBR Green dye uptake assays in terms of specificity.

During the course of the NSN submissions, Environment Canada requested that DRDC – Suffield Research Centre use a definitive method to confirm that the *Escherichia coli* host strain for MS2 propagation (*E. coli* K-12S A/λ (F+)) was indeed a K-12 strain. A gel-based PCR assay from the open literature that was designed to detect a common K-12 genetic marker was modified and used to test the DRDC – Suffield Research Centre *E. coli* host strain, along with a number of negative and positive control strains from the DRDC – Suffield Research Centre collection. In addition, *in silico* analysis was conducted to address Environment Canada's concern regarding the specificity of the *aro*Q assay to detect the DRDC – Suffield Research Centre simulant *Pantoea agglomerans* ATCC 33243, but not related *Pantoea* strains, including plant pathogens.

This paper describes the scientific methods and analysis that was conducted in support of DRDC – Suffield Research Centre's NSN submissions of biological agent simulants to Environment Canada.

2 Materials and methods

2.1 Bacterial and viral cultures

The bacterial and viral culture materials used in this project are presented in Table 1.

Table 1: *List of bacterial and viral materials used.*

Item	Organism	Material Tested	Purpose	K12 Reference
1	MS2 bacteriophage (ATCC 15597-B1)	cell-free lysate	assay positive control	
2	*Escherichia coli* K-12S A/λ (F+)	cell lysate	test material	
3	*Pantoea agglomerans* (ATCC 33243)	cell lysate	assay positive control	
4	*Bacillus atrophaeus*	cell lysate	assay positive control	
5	*E. coli* K12 strain ATCC 15597	cell lysate	K-12 control strain	[21]
6	*E. coli* K12 strain BMH71-18mutL	cell lysate	K-12 control strain	[24]
7	*E. coli* K12 strain DH5α	cell lysate	K-12 control strain	[23]
8	*E. coli* K12 strain JM101	cell lysate	K-12 control strain	[25]
9	*E. coli* K12 strain JM109	cell lysate	K-12 control strain	[25]
10	*E. coli* K12 strain TG2	cell lysate	K-12 control strain	[22]
11	*Acinetobacter sp.* ATCC 49139	gDNA	non K-12 control strain	
12	*Enterobacter aerogenes* ATCC 13048	gDNA	non K-12 control strain	
13	*Pantoea agglomerans* ATCC 33243	gDNA	non K-12 control strain	
14	*Proteus vulgaris* ATCC 8427	gDNA	non K-12 control strain	
15	*Klebsiella pneumoniae* ATCC 9997	gDNA	non K-12 control strain	
16	*Pseudomonas aeruginosa* ATCC 35032	gDNA	non K-12 control strain	

Fresh cultures of items 1 through 5 were prepared in-house. MS2 bacteriophage (Item 1) was prepared to a concentration of 5.3×10^{10} PFU/mL in Luria broth media. The original MS2 culture was obtained from Dugway Proving Grounds (DPG, West Desert Test Center, Dugway, UT); this was later confirmed as ATCC 15597-B1. A culture plate of pure *E. coli* K-12S A/λ (F+) was included (Item 2); this strain was used as the MS2 host in previous field trials and also originated from DPG. *Pantoea agglomerans* (Item 3), *Bacillus atrophaeus* NSN #18250-7 (Item 4) and the K-12 control strain *E. coli* K12 ATCC 15597 (Item 5) [21] were propagated on agar plates. Cultures were stored at 4°C prior to preparing cell lysates for PCR analysis.

Items 6–10 were *E. coli* K-12 control strains that were prepared from freeze-dried vials held in the DRDC – Suffield Research Centre strain collection and stored at 4°C [22–25]. The vial contents were hydrated with 100 µL of sterile phosphate-buffered saline (PBS). A 10 µL sterile, disposable loop was used to aseptically plate hydrated cells onto a Luria-Bertani (LB) agar plate media prepared in house (Sigma-Aldrich Canada Ltd., Oakville, ON) and incubated at 37°C for 24 to 48 h. The plates were removed from the incubator, wrapped in parafilm, and stored at 4°C prior to preparing cell lysates for PCR analysis.

Items 11–16 were genomic DNA (gDNA) preparations isolated from non *E. coli* strains in the DRDC – Suffield Research Centre strain collection using the Qiagen EZ1 instrument according to the manufacturer's instructions (Qiagen, Mississauga, ON). The DNA was quantified using the Nanodrop spectrophotometer according to the manufacturer's instructions (NanoDrop Technologies, Inc., Wilmington, DE) and adjusted to a working concentration of 0.5 ng/µL with 0.2 µm, filter-sterilized TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0).

2.2 Soil isolates

Isolates from nine EPG soils processed for bacterial spore recovery and demonstrating vegetative growth on agar plates were provided on agar culture plates (labelled S1, S2, S3, S7, S8, S9, S10, S11 and S12) for PCR analysis using the *Bacillus atrophaeus* assay. The plates were stored at 4°C prior to preparing cell lysates for PCR analysis.

2.3 Preparation of cell-free MS2 virus positive control template—Boil Prep method

The stock MS2/*E. coli* ATCC 15597-B1 culture (~5.3 × 10^{10} PFU/mL) was diluted 10-fold to a total volume of 2 mL (1.8 mL of 10 mM Tris, 1 mM EDTA, pH 7.5 buffer + 0.2 mL of MS2/*E. coli* culture) and drawn into a 5 mL syringe. A 0.22 µm nylon syringe filter (0.22 µm × 25 mm) was screwed onto the syringe and the MS2/*E. coli* culture was filtered into a sterile 2.0 mL Simport microfuge tube with o-ring screw cap (Mandel Scientific, Guelph, ON) to remove *E. coli* cells (a 20% $^{v}/_{v}$ loss in recovery was observed due to filter absorption). All Simport tubes used in this project were RNase-, DNase-, Pyrogen-, and DNA-free. The tube contents were spun down briefly (1–2 sec, microspinner full speed). A 1 mL aliquot of the filtered MS2 supernatant was transferred into a clean 2 mL microfuge tube with o-ring screw cap and immersed into a boiling water bath (~ 100°C) for 10 minutes to prepare a boiled lysate. The tube was removed from the boiling water bath and spun down briefly (1–2 sec; microspinner full speed). The contents were pipeted into 10 × 100 µL volumes. The tubes were stored at −70°C (4 tubes), 4°C (4 tubes) and room temperature (RT) (2 tubes). One of the tubes was used to make a serial dilution series for a standard curve for the MS2 assay 4 and assay 5 using DNase-free, RNase-free water (NFW) from Applied Biosystems® (Streetsville, ON) as the diluent.

2.4 Preparation of cell-free MS2 virus positive control template—TRIzol® method

A 1 mL aliquot of filtered MS2 supernatant was prepared as described above. A 20 µL aliquot of filtered MS2 (~5.3x10^{9} PFU/mL) was added to 500 µL of TRIzol® (Life Technologies-Invitrogen, Burlington, ON).[2] The solution was mixed by pipet and incubated at RT for 5 min.[3] One hundred microlitres of chloroform was added to the TRIzol®/virus tube. The tube was capped and shaken vigorously by hand for 15 sec and incubated at RT for 2–3 min. The tube was centrifuged at 12000 × g for 15 min at 4°C to separate the phases. The top aqueous phase containing the RNA (~250 µL) was transferred using a 1 mL pipet to a clean tube. The RNA was precipitated by adding 250 µL of isopropanol, mixing by inversion, and incubating at RT for 10 min. The tube was centrifuged at 12000 × g for 10 min at 4°C to pellet the RNA. The supernatant was removed to waste by pipet and the pellet was washed with 0.5 mL of 75% DEPC-treated ethanol.[4] The pellet was washed by mixing on the vortex followed by centrifugation at ≤ 7500 × g for 5 min at 4°C. The pellet was air-dried for 5–10 min making sure not to let it dry to completion. The pellet was dissolved in 50 µL of NFW and incubated at 55–60°C for 10 min to help dissolve the RNA pellet. The RNA solution was measured by spectroscopy on the Nanodrop (OD 260/280) and then aliquots were stored at −70°C, 4°C, and ambient temperature. Prior to use, an

[2] Work with TRIzol® reagent was conducted in the fumehood while wearing gloves and eye protection.
[3] Samples can be stored at this point at −70°C for at least one month.
[4] This can be stored at 4°C for at least one week and at −20°C for at least 1 year.

aliquot was removed from storage and a 10-fold serial dilution series was made for a standard curve for the MS2 assay 4 and assay 5 using NFW as the diluent.

2.5 Preparation of bacterial cell lysates—Boil Prep method

Individual colonies or cell lawns from pure cultures grown on agar plates were selected using a sterile disposable loop and suspended in NFW in Simport microfuge tubes. The cells were mixed by pipet using sterile, plugged pipet tips or by vortex at maximum speed for ~5 seconds. After mixing, tubes were spun down briefly (1–2 sec) on a microspinner (full speed). The tubes were then immersed in a boiling water bath (~100°C) for 10 min. The tubes were removed from the water bath and spun at 18,000 × rcf for 30 seconds (Beckman Coulter Microfuge® 18 Centrifuge). The supernatant containing the template was removed by pipet to a new microfuge tube. The tube containing the pellet was discarded to waste. Serial dilutions of the supernatant were prepared using NFW. Neat and diluted cell lysate supernatants were stored at 4°C or −20°C prior to analysis. For *Bacillus atrophaeus* and *Pantoea agglomerans*, eight colonies were suspended in 100 µL of NFW. For *E. coli* ATCC 15597, a single colony was suspended in 50 µL of NFW. All other *E. coli* K-12 strains were prepared by scraping a loopful of cells using a disposable, sterile 10 µL plastic loop from purified lawns previously streaked and shown to produce individual colonies of identical phenotype (size, shape, and color) on LB agar plates. The cells were suspended in 200 µL of NFW prior to boiling.

2.6 PCR reaction setup and run parameters for the simulant assays

The PCR assays used in this project for the simulants MS2 [15], *B. atrophaeus* [26], and *P. agglomerans* [27] were derived from the open literature. A concentrated PCR master mix at 1.087× strength was made for a given assay in bulk volume (# reactions × 25 µL × 20% extra) in a template-free biosafety cabinet (BSC) using commercially available Taqman® master mix kits from Life Technologies Inc, Applied Biosystems® (Burlington, ON) (Table 2). Template-free PCR reactions (also referred to as no template controls or NTCs) were prepared in the same template-free BSC by adding 6.52 µL of NFW to 75 µL of 1.087× PCR master mix to prepare a 1× strength PCR mixture. The concentration of the reaction components in the final PCR reaction volume of 25 µL are listed in Table 2 (where known).

Table 2: *Concentration of reaction components in the final PCR reaction.*

	B. atrophaeus **Assay [final]**	***P. agglomerans*** **Assay [final]**	**MS2 Assay [final]**
Master Mix Components	Taqman® PCR Core Reagents Kit	Taqman® PCR Core Reagents Kit	Taqman® One-Step RT-PCR Master Mix Reagents Kit
Taqman® buffer A	1×	1×	Not applicable
MultiScribe™ reverse transcriptase/ RNase Inhibitor	Not applicable	Not applicable	1×
$MgCl_2$	5 mM	5 mM	unknown[a]
dUTP	0.2 mM	0.2 mM	unknown[a]
dATP	0.1 mM	0.1 mM	unknown[a]
dCTP	0.1 mM	0.1 mM	unknown[a]
dGTP	0.1 mM	0.1 mM	unknown[a]
AmpliTaq Gold® DNA polymerase	0.05 U/µL	0.05 U/µL	unknown[a]
AmpErase® uracyl N-glycosylase	0.01 U/µL	0.01 U/µL	unknown[a]
Forward Primer	0.2 µM	0.3 µM	0.4 µM
Reverse Primer	0.2 µM	0.3 µM	0.4 µM
Probe	0.2 µM	0.2 µM	0.2 µM
ROX reference dye	0.016 µM	0.016 µM	0.016 µM[b]

[a] Concentration or presence in the mixture was not provided by the manufacturer.

[b] Also contains passive reference 1 in manufacturer's supplied buffer.

A 25 µL aliquot of 1× PCR mixture was pipetted into a well in a 96 well optical plate (Applied Biosystems®). This was repeated two more times for a total of three replicates per sample. The plate was placed onto a 96 well cold block and then transferred from the template-free BSC to a second BSC for the addition of test material (samples and positive controls) by adding 6.52 µL of test material to 75 µL of 1.087× PCR master mix to prepare a 1× strength PCR mixture. A 25 µL aliquot of 1× PCR mixture was pipetted into a well in the 96 well optical plate and repeated two more times for a total of three replicates per sample. The plate was sealed with Microamp optical adhesive film (Applied Biosystems®). The plate contents in the sealed plate were spun down for approximately two minutes at 1500–2000 rpm (Beckman Allegra™ 6R centrifuge). The plate was loaded onto the 7500 FAST Cycler unit (Applied Biosystems®) and the run was setup following the manufacturer's guide for setting up a Relative Quantification (RQ) experiment.

2.6.1 Probes and primers for the simulant assays

All probes and primers used for the simulant assays were prepared by Applied Biosystems® – Life Technologies Corporation, (Carlsbad, CA). All Taqman probes were fluorescently labelled on the 5′ end with 6-FAM (6-Carboxyfluorescein) and on the 3′ end with TAMRA (Carboxytetramethylrhodamine).

2.6.2 PCR cycling parameters for the simulant assays

The PCR cycling programs were run according to the following parameters:

Cycling Program for the *Bacillus atrophaeus* and *Pantoea agglomerans* Assays

50°C for 2 min (1 cycle)
95°C for 10 min (1 cycle)
95°C for 15 sec + 60°C for 60 sec (40 cycles)

Cycling Program for the MS2 Assay

48°C for 30 min (1 cycle)
95°C for 10 min (1 cycle)
95°C for 15 sec + 60°C for 60 sec (45 cycles)

The sequences of the primers and probes for each assay are listed in Table 3.

Table 3: *Simulant PCR assay primer and probe sequences.*

Assay (Gene target)	Type	Sequence (5′ to 3′)	Reference
B. atrophaeus (*rec*A gene)	Primer 1	acc-aga-caa-tgc-tcg-acg-tt	[27]
	Primer 2	ccc-tct-tga-aat-tcc-cga-at	
	Probe	FAM-act-gaa-cag-ctg-atc-gag-aca-gct-gca-TAMRA	
P. agglomerans (*aro*Q gene)	Primer 1	gct-gca-aaa-cgc-aca-aca	[26]
	Primer 2	cgt-gaa-caa-acg-gct-cca	
	Probe	FAM-ccg-ggc-ttg-aac-ccc-act-cc -TAMRA	
MS2 assay 4 (Lysis protein)	Primer 1	cct-cag-caa-tcg-cag-caa-a	[15]
	Primer 2	gga-aga-tca-ata-cat-aaa-gag-ttg-aac-ttc	
	Probe	FAM-caa-aca-tga-gga-tta-ccc-atg-tcg-aag-aca-TAMRA	
MS2 assay 5 (RNA replicase β chain)	Primer 1	gct-ctg-aga-gcg-gct-cta-ttg	[15]
	Primer 2	cgt-tat-agc-gga-ccg-cgt	
	Probe	FAM-ccg-aga-cca-atg-tgc-gcc-gtg-TAMRA	

The analysis settings used to generate the PCR data for the simulant assays on the 7500 FAST platform are presented in Table 4.

Table 4: *Real-Time PCR assay 7500 FAST analysis settings.*

Item	Setting
Threshold	Set above the background within the exponential phase of the amplification curve and is recorded during the Ct data collection and reporting functions by the instrument software
Calibrator sample	NTC (no template control)
Endogenous control	None
Control Type	None
RQ Min/Max (basic)	1.0
Auto Ct	Manual
Baseline	3,15

2.7 PCR reaction setup and run parameters for the *E. coli* K12 assay

The PCR assay used for assessing DRDC – Suffield Research Centre's MS2 host *E. coli* strain K-12S A/λ (F+) as a K-12 strain, was obtained from the published literature [23] but with some modifications. Briefly, PCR reactions were made as a 1.087× concentrate (23 µL) and brought to 1× PCR strength by adding 2 µL of lysate or 2 µL of NFW in the case of the NTC reactions. The master mix was composed of buffer, $MgCl_2$, dNTPs, and Taq DNA polymerase from the LightCycler® 480 Probes Master mix supplied at 2× strength (Roche Canada, Mississauga, ON). The primers were obtained from IDT (Coralville, IA). The final concentrations of the PCR reaction components are provided in Table 5 (where known). Primer target and sequence details are provided in Table 6.

Table 5: *Concentration of reaction components in the E. coli K12 assay PCR reaction.*

	***E. coli K12 assay* [final]**
Master Mix Components	LightCycler® 480 Probes Master mix
Reaction buffer	1×
$MgCl_2$	3.2 mM
dUTP	?
dATP	?
dCTP	?
dGTP	?
FastStart Taq DNA Polymerase	?
Forward Primer	0.4 µM
Reverse Primer	0.4 µM

Table 6: *E. coli K12 PCR primer sequences.*

Gene target	Primer Type (Name)	Sequence (5′ to 3′)	Amplicon size	Reference
orf 264 region of the *E. coli* K-12 genome	Forward Primer (K12-R)	atc-ctg-cgc-acc-aat-caa-caa	969 bp	[23]
	Reverse Primer (K12IS-L)	cgc-gat-gga-aga-tgc-tct-gta		

Two NTC reactions were prepared and run in triplicate. NTC1 was prepared in a template-free area. The NTC1 tubes were then closed and remained closed until loaded on the gel. NTC2 was prepared in a template-free area, closed and moved to the template-addition area, where the tubes were opened and left open until the addition of template was completed. The NTC2 tubes were then closed and remained closed until loaded on the gel. The cell lysate of *E. coli* strain K-12S A/λ (F+), prepared in triplicate was the first template-containing sample to be prepared to ensure that this sample could not be contaminated with positive control template from the remaining positive control strains. The *E. coli* positive control PCR reactions were subsequently prepared next in the following order: ATCC 15597, BMH71-18mutL, DH5α, JM101, JM109, and TG2. Non-*E. coli* PCR reactions were prepared subsequently and included gDNA from *Acinetobacter sp*. ATCC 49139, *Enterobacter aerogenes* ATCC 13048, *Pantoea agglomerans* ATCC 33243, *Proteus vulgaris* ATCC 8427, *Klebsiella pneumoniae* ATCC 9997, and *Pseudomonas aeruginosa* ATCC 35032.

PCR was performed on the Eppendorf Mastercycler® Gradient (Mississauga, ON) using 0.5 mL thin-walled Eppendorf PCR tubes. The PCR cycling programs were run according to the following parameters:

94°C for 10 min (1 cycle)
94°C for 30 sec + 60°C for 30 sec + 72°C for 180 sec (35 cycles)

2.7.1 Gel analysis

Gel analysis was used to visualize the PCR amplification reaction products for the *E. coli* K-12 PCR assay. Five µL of each PCR reaction were added to 5 µL of 2× strength gel loading buffer (final concentration was 0.01% bromophenol blue, 0.1% SDS, 0.01M EDTA, 5% glycerol), mixed, and then loaded onto a 1.2% agarose gel (UltraPure™ Agarose, Invitrogen Life Technologies, Burlington, ON) containing 0.3 mg/mL ethidium bromide along with molecular weight markers (Track-iT 100 bp DNA ladder markers, Invitrogen, Burlington, ON). The gel was electrophoresed in 1× TBE running buffer (0.09M Tris-base, 0.09M boric acid, 0.002M EDTA, pH 8.3) for 105 minutes at 112 volts. Gel results were captured using the SynGene Geni gel imager (Discovery Scientific Inc., Vancouver, BC).

3 Results

3.1 Selection of an MS2 assay and an MS2 positive control preparation method

An evaluation of five real-time, probe hydrolysis, reverse-transcription assays for MS2 detection [15] resulted in two assays (MS2 assay 4 and assay 5) exhibiting the lowest level of detection among the five assays (0.4 fg or about 200 genomic equivalents). Both of these assays were selected for evaluation in this study using MS2 positive control template prepared by two different techniques, namely, the TRIzol® extraction method and the Boil Prep method. Real-time, reverse-transcription PCR for both MS2 assays against serially diluted MS2 control template indicated that MS2 assay 5 was more sensitive than MS2 assay 4 for both template preparation methods (Annex A). Thus MS2 assay 5 was employed for the remainder of this study. Because the Boil Prep method for preparing the MS2 template was simple, quick, and effective, the boil treatment was used to prepare the positive control template for the bacterial templates as well.

3.2 PCR assay standard curves for simulant assays using template prepared by the boiling prep method

Positive control template (PC) for each of the three simulants prepared using the Boil Prep method was serially diluted and tested to determine a suitable working dilution to use for testing unknown isolates (Annex B). Lines of best fit were obtained for each assay across a dilution series of their respective positive control. A working dilution of 1/10 for each positive control was selected as a suitable dilution for testing. An expected Ct value for each positive control at 1/10 dilution was calculated from the line of best fit and is presented in Table 7.

Table 7: *Expected Ct values for 1/10 dilutions of positive controls from standard curves.*

	***B. atrophaeus* Assay**	***P. agglomerans* Assay**	**MS2 Assay**
Line of best fit equation	$y = -1.6 \ln(x) + 16.798$	$y = -1.609 \ln(x) + 11.142$	$y = -1.446 \ln(x) + 14.403$
R^2 value	0.9994	0.9983	0.9899
Ct value at 1/10 dilution	20.48	14.8	17.73
Threshold setting	0.0992875	0.0992875	0.0670019

3.3 Cross-reactivity testing of simulant assays

Cross-reactivity testing of each simulant assay was performed to ensure that a given simulant assay did not cross-react with the DNA from the other simulants being tested. Each assay was tested against 1/10 dilutions of positive control material from all the simulants. Each assay was shown to be specific for its intended target and did not cross-react with positive control material from heterologous simulants (Table 8).

Table 8: *Evaluation of simulant assays against heterologous positive controls.*

	Average Ct value ± Std Dev (n = 3)			
	***B. atrophaeus* PC**	***P. agglomerans* PC**	**MS2 PC**	**Threshold setting**
B. atrophaeus assay	20.639 ± 0.088	0.00	0.00	0.0992875
P. agglomerans assay	0.00	15.762 ± 0.033	0.00	0.0992875
MS2 assay	0.00	0.00	17.114 ± 0.029	0.0670019

3.4 PCR analysis of spore isolates from DRDC – Suffield Research Centre field soils

Nine culture plates (S1–S3, S7–S9 and S10–S12), each containing a pure culture of vegetative cells isolated from EPG soils using the spore selection method, were analyzed by PCR as these were the only cultures isolated. Since these cultures were propagated using the spore selection method, they were tested using the *B. atrophaeus* PCR assay as *B. atrophaeus* was the only simulant of the three in this study that forms spores. The Boil Prep method was used to prepare material from selected colonies on each plate and then tested by PCR on the 7500 FAST platform (Threshold = 0.0992875). The results are presented in Table 9.

Table 9: *Analysis of EPG soil isolates using the Bacillus atrophaeus PCR assay.*

		Plate	**Ct 1**	**Ct 2**	**Ct 3**	**Avg Ct**	**Ct Std Dev**
CTTC Soil Isolate Test Samples[a]	S1–S3	1	21.096	21.041	20.964	21.034	0.066
		2	21.351	21.497	21.503	21.450	0.086
		3	21.381	21.601	21.578	21.520	0.121
	S7–S9	1	21.286	21.215	21.227	21.243	0.038
		2	20.520	20.454	20.333	20.436	0.095
		3	20.721	20.792	20.975	20.829	0.131
	S10–S12	1	21.129	21.047	21.169	21.115	0.062
		2	20.622	20.621	20.596	20.613	0.015
		3	20.375	20.267	20.210	20.284	0.084
PC[b]	BG PC	NA	20.549	20.724	20.646	20.639	0.088
	EH PC	NA	0.000	0.000	0.000	0.000	0.000
	MS2 PC	NA	0.000	0.000	0.000	0.000	0.000
NTC		NA	0.000	0.000	0.000	0.000	0.000

[a] 1/10 dilutions of boil preparations derived from single colonies (prepared 4 Oct 2010).

[b] Positive controls prepared as 1/10 dilutions of boil preparations of BG, EH, and MS2 respectively.

Boil preps from all purified isolates derived from each of the three soil samples (S1–S3, S7–S9, and S10–S12) were positive using the *B. atrophaeus* PCR assay. The only positive control that was PCR-positive was the *B. atrophaeus* positive control, as expected. The no-template control (NTC) was PCR-negative, as expected. Thus all nine isolates were PCR-confirmed as *B. atrophaeus*.

3.5 PCR analysis of boil preparations of purified colonies obtained from "Old BG" (circa 1963) and "New BG" (circa 1996) DRDC – Suffield Research Centre stocks

Boiled lysates from replicate colonies (three colonies per sample) were prepared from plate cultures of "Old BG" and "New BG" and analyzed using the *B. atrophaeus* PCR assay on the 7500 FAST platform using a threshold setting of 0.0992875 (Table 10) along with positive and negative controls.

Table 10: *PCR analysis of colony lysates of "Old BG" and "New BG" using the Bacillus atrophaeus PCR assay.*

		colony	Ct 1	Ct 2	Ct 3	Avg Ct[c]	Ct Std Dev
Test Samples[a]	Old BG	1	21.061	21.012	20.898	20.990	0.084
		2	20.444	20.604	20.564	20.537	0.083
		3	20.764	20.886	20.922	20.857	0.083
	New BG	1	20.630	20.708	20.726	20.688	0.051
		2	21.214	21.158	21.058	21.143	0.079
		3	20.568	20.557	20.743	20.621	0.105
PC[b]	BG PC	NA	20.788	20.776	20.748	20.771	0.021
	EH PC	NA	0.000	0.000	0.000	0.000	0.000
	MS2 PC	NA	0.000	0.000	0.000	0.000	0.000
NTC		NA	0.000	0.000	0.000	0.000	0.000

[a] 1/10 dilutions of boiled lysates derived from single colonies of "Old BG (circa 1963) and "New" BG (circa 1996).

[b] Boiled lysates of BG and EH (1/10 dilutions) and MS2 (1/1000 dilution).

[c] Ct data based on a threshold setting of 0.0992875.

The *B. atrophaeus* assay generated PCR-positive signals of equivalent response for both the "Old BG" and "New BG" cultures for all colonies tested.

3.6 PCR confirmation of DRDC – Suffield Research Centre MS2 host *E. coli* K-12S A/λ (F+) strain as a K-12 strain

Following the NSN submission of the MS2 host *Escherichia coli* strain K-12S A/λ (F+) to Environment Canada, DRDC – Suffield Research Centre was asked to confirm that it was a K-12 strain, using a more definitive method than that which had been previously submitted. Failure to achieve this would have resulted in termination of review of the NSN package, thereby requiring re-submission and substantial additional time and effort that would have prevented planned field use of this simulant by DRDC – Suffield Research Centre in the fall of 2011. Given the short timeline (~one month) to propose a suitable test method and then conduct the analysis, DRDC – Suffield Research Centre proposed using a published gel-based PCR assay specific for identifying *E. coli* K-12 strains [23], which Environment Canada accepted.

Because the DRDC – Suffield Research Centre strain collection contained a number of *E. coli* K-12 strains and gDNA from a number of non-*E. coli* strains to use as positive and negative PCR controls, respectively, procurement of test materials was not required, other than the procurement of PCR primers and reagents. Cell lysates from the DRDC – Suffield Research Centre *E. coli* K-12S A/λ (F+) strain and several *E. coli* K-12 strains in the DRDC – Suffield Research Centre strain collection were prepared and amplified by PCR along with gDNA from a number of non-*E. coli* bacterial strains. The PCR reaction products were separated and visualized by horizontal agarose gel analysis (Figure 1).[5]

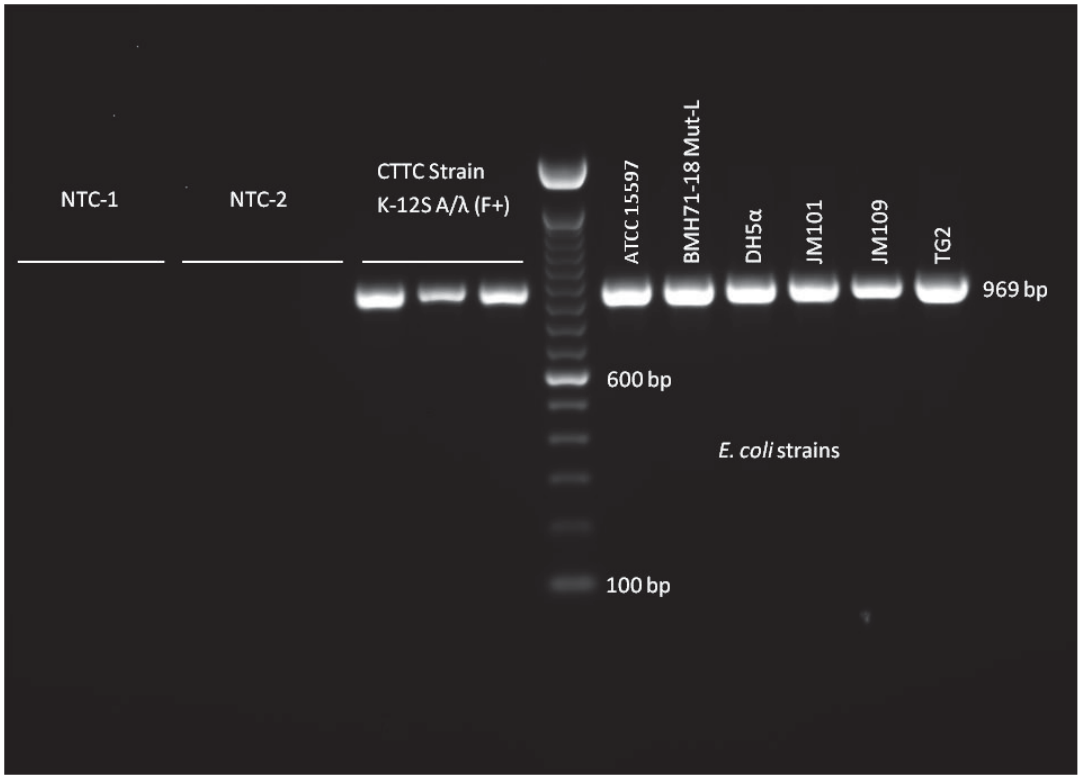

Figure 1: *Gel image of PCR amplified lysates from various E. coli K-12 strains and the DRDC – Suffield Research Centre E. coli test strain K-12S A/γ (F+).*

All three replicate cell lysates of the DRDC – Suffield Research Centre *E.coli* K-12S A/λ (F+) strain generated a PCR amplicon of the size expected for K-12 strains (~969 bp fragment in Figure 1). An amplicon band of the same size and intensity as the DRDC – Suffield Research Centre test strain was observed for each of the *E. coli* K-12 positive control strain lysates as well. No amplicons were observed for the non-*E. coli* control strains which included *Acinetobacter sp*. ATCC 49139, *Enterobacter aerogenes* ATCC 13048, *Pantoea agglomerans* ATCC 33243, *Proteus vulgaris* ATCC 8427, *Klebsiella pneumoniae* ATCC 9997, and *Pseudomonas aeruginosa* ATCC 35032 (results not shown). Based on these results, the DRDC – Suffield Research Centre *E. coli* K-12S A/λ (F+) strain used as a host for the field simulant MS2 is a K-12 strain. This data was submitted in time for the September review and was accepted by Environment Canada, thereby allowing the field trials at DRDC – Suffield Research Centre to proceed without delay.

[5] Samples were electrophoresed in 1×TBE buffer through a 1.2% agarose gel containing 0.3 mg/mL ethidium bromide for 105 minutes at 112 volts. Image was captured using the SynGene Geni gel imager (0.423 s exposure).

3.7 Analysis of published data and information concerning CTTC's *P. agglomerans* strain (ATCC 33243)

Due to the lack of published information on *P. agglomerans* ATCC 33243, DRDC – Suffield Research Centre's NSN submission package used strain E325 as a surrogate for some of the information. In order to accept the E325 surrogate data, Environment Canada requested molecular genetic evidence that ATCC 33243 strain was similar to the commercially available biological control agent E325 [6]. ATCC 33243 and E325 were considered to have reasonable similarity based on BLAST analysis of partial 16S sequence data performed by Dr. Chris Wend, Director of Bioscience Division of Northwest Agricultural Products (Pasco, WA) for DRDC in June 2009 (Annex C). Environment Canada requested that the raw 16S sequences for these two strains be provided, as well as that of *P. agglomerans* C9-1, which is also a biocontrol strain [5]. The required 16s rRNA sequence data was downloaded and compiled (Annex D) to meet this request.

In addition, Environment Canada asked whether the *aro*Q assay used by DRDC – Suffield Research Centre for the confirmation of *P. agglomerans* ATCC 33243 was strain-specific or if it could differentiate *P. agglomerans* ATCC 33243 from other *P. agglomerans* strains, particularly strains found to be plant pathogens. To answer this, an NCBI BLAST analysis (blastn) of the *aro*Q assay primers against the NR Genbank sequence database and the *Pantoea* Taxid (taxid 53335) was conducted. *P. agglomerans* ATCC 33243 (Genbank M95628.1) was the only strain recognized by both the forward and reverse primers across the entire primer sequence (18/18 matches for the forward primer and 18/18 matches for the reverse primer), suggesting the assay to be selective for *P. agglomerans* ATCC 33243. Three other *Pantoea* species were reported but these were based on partial alignments and only to the forward primer: *Pantoea ananatis* AJ13355 complete genome (14 of 18 matches); *Pantoea ananatis* LMG 20103 complete genome (14 of 18 matches); and *Pantoea vagans* C9-1 complete genome (14 of 18 matches). Interestingly, *Pantoea ananatis* LMG 20103 is a plant pathogen, while *Pantoea vagans* C9-1, formerly *Pantoea agglomerans* C9-1, is a biocontrol strain [28]. Surprisingly, the E325 biocontrol strain was not picked up during the BLAST analysis. This likely occurred as a result of a lack of sequence data for the *aro*Q gene of strain E325 in the database.

When the amplified gene sequence of the *P. agglomerans* ATCC 33243 strain was selected for NCBI BLAST analysis (blastn) against the NR database and all bacterial sequences, the *P. agglomerans* ATCC 33243 amplicon sequence (M95628.1) matched against itself, as expected, with a maximum identity of 100% across the entire queried sequence, with a maximum score of 107, and an E-value of 4×10^{-21}. The next closest match was Genbank sequence CP001893.1, which is a gene sequence on the plasmid pPag1 of *Pantoea vagans* C9-1 (biocontrol strain), with a maximum identity of 80%, a maximum score of 53.6, and an E. value of $7x10^{-5}$. Sequence alignment between these two strains across the amplicon region (59 nucleotides) revealed two mismatches in the forward primer region, three mismatches in the reverse primer region, and five mismatches in the probe region (Figure 2).

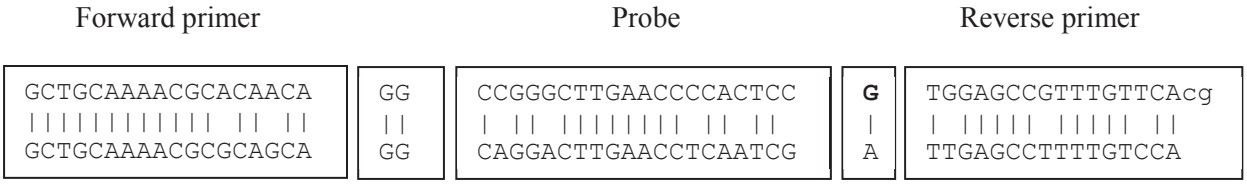

Figure 2: *Sequence alignment of the aroQ amplicon of the DRDC – Suffield Research Centre simulant Pantoea agglomerans ATCC 33243 (top) and the biocontrol strain Pantoea vagans C9-1 (bottom).*

Thus, based on in silico analysis, the *aro*Q assay used in this study will detect the DRDC – Suffield Research Centre simulant *P. agglomerans* ATCC 33243, but should not be able to detect the biocontrol strain *Pantoea vagans* C9-1 or plant pathogen strains such as *Pantoea ananatis* LMG 20103. Lab-testing to confirm this was not requested by Environment Canada and thus was not pursued in this study, but testing the assay against these strains and other control strains would address any uncertainty.

4 Discussion

The PCR simulant assays described in this publication were adapted for use at DRDC – Suffield Research Centre to confirm isolates derived from environmental samples in support of its New Substances Notification submissions to Environment Canada and to confirm the identity of these simulants in future applications for both field and laboratory use. Because *B. atrophaeus* was the only simulant of the three to be recovered from the DRDC – Suffield Research Centre EPG site survey of soils, the *B. atrophaeus* assay was the only assay that was used to confirm the identity of the soil isolates.

The assessment of two PCR assays for MS2 revealed one of the assays to be slightly better due to lower Ct values at equivalent dilutions. The Boil Prep method was found to be easier and safer to use than the TRIzol® method for preparing PCR amplifiable template, since the TRIzol® method employs harsh chemicals, requires several steps, and requires one to work in the fumehood and wear protective equipment to minimize potential exposure to the TRIzol® reagent. Thus the Boil Prep method was selected as the method of choice to prepare control template material for this project. Confirmation of DRDC – Suffield Research Centre's *E. coli* host strain as a K-12 strain using an *E. coli* K12-specific PCR assay was an important component of the NSN submission for MS2. The ability to use the *E. coli* K-12 strains and non-*E. coli* strain DNA from the DRDC – Suffield Research Centre strain collection as positive and negative PCR controls, respectively, was critical in confirming the specificity of the published assay and critical in using this assay in a timely fashion, thereby allowing DRDC – Suffield Research Centre to use this material in the fall of 2011 as planned. Furthermore, failure to meet Environment Canada's deadline would have resulted in termination of review of the NSN package, thereby requiring re-submission and substantial additional time and effort.

During this project, two different batches of BG, namely an old batch prepared by Dugway Proving Ground *circa* 1963 and a new batch prepared by Dugway *circa* 1996 were tested. The *B. atrophaeus* PCR assay was positive for replicates taken from both batches with essentially no difference in Ct values. Thus, these two batches likely contained the same genetic targets detected by the *B. atrophaeus* assay. Whole genome sequencing of isolates from both batches would be able to distinguish genetic differences across the entire genome if they existed, but the time and cost to do this was not warranted for this project.

In silico analysis was conducted to address Environment Canada's concern regarding the specificity of the *aro*Q assay to detect the DRDC – Suffield Research Centre simulant *Pantoea agglomerans* ATCC 33243, but not related *Pantoea* strains, including plant pathogens. *In silico* analysis indicated that the *aro*Q assay should detect the DRDC – Suffield Research Centre simulant strain but not the biocontrol strain *Pantoea vagans* C9-1 or plant pathogen strains such as *Pantoea ananatis* LMG 20103. Lab-testing to confirm this was not requested by Environment Canada and thus was not pursued in this project, but testing the assay against these strains and other control strains would address any uncertainty.

The data and information presented in this report contributed to the NSN submissions to Environment Canada and resulted in the approval to use MS2/*E.coli* and *Pantoea agglomerans* for outdoor aerosol

release at the DRDC – Suffield Research Centre field site as described in the NSN application.[6] *Bacillus atrophaeus* (BG) was approved for outdoor use as well, but was not required to go through the NSN submission process, as it was subsequently placed on the NSN Program Domestic Substances List due to DRDC – Suffield Research Centre's historical use. All of the assays used in this project can be used in future applications for both field and laboratory use.

[6] To be published by DRDC Suffield separately.

5 Conclusion

PCR assays for biothreat simulant organisms that are used by DRDC – Suffield Research Centre for training and equipment assessment purposes were adapted to confirm the identity of simulant strains and to support its New Substances Notification submissions to Environment Canada. The data and information presented in this report resulted in the approval to use MS2/*E.coli* and *Pantoea agglomerans* for outdoor aerosol release at the DRDC – Suffield Research Centre field site as described in the NSN application. The *B. atrophaeus* (BG) assay was used to confirm cultured spore isolates recovered from soils collected from the DRDC – Suffield Research Centre Experimental Proving Ground as *B. atrophaeus*. The assay was also used to analyze two different batches of BG (originally prepared *circa* 1963 and *circa* 1996). *Bacillus atrophaeus* was approved for outdoor use by Environment Canada, but was not required to go through the NSN submission process, as it was subsequently placed on the NSN Program Domestic Substances List due to DRDC DRDC – Suffield Research Centre's historical use. All of the assays used in this project can be used in future applications for both field and laboratory use.

References

[1] Cruz, A.T., A.C. Cazacu, and C.H. Allen, *Pantoea agglomerans,* a plant pathogen causing human disease. *Journal of Clinical Microbiology*, 2007. **45**(6): pp. 1989–1992.

[2] Barash, I. and S. Manulis-Sasson, Recent evolution of bacterial pathogens: the gall-forming *Pantoea agglomerans* case. *Annual Review of Phytopathology*, 2009. **47**: pp. 133–152.

[3] Weinthal, D.M., et al., Distribution and replication of the pathogenicity plasmid pPATH in diverse populations of the gall-forming bacterium *Pantoea agglomerans*. *Applied and Environmental Microbiology*, 2007. **73**(23): pp. 7552–7561.

[4] De Champs, C., et al., Isolation of *Pantoea agglomerans* in two cases of septic monoarthritis after plant thorn and wood sliver injuries. *Journal of Clinical Microbiology*, 2000. **38**(1): pp. 460–461.

[5] Rezzonico, F., et al., Genotypic comparison of *Pantoea agglomerans* plant and clinical strains. *BMC Microbiology*, 2009. **9**(204).

[6] Pusey, P.L., V.O. Stockwell, and D.R. Rudell, Antibiosis and acidification by *Pantoea agglomerans* strain E325 may contribute to suppression of *Erwinia amylovora*. *Phytopathology*, 2008. **98**(10): pp. 1136–1143.

[7] Volksch, B., et al., Polyphasic study of plant- and clinic-associated *Pantoea agglomerans* strains reveals indistinguishable virulence potential. *Infection Genetics and Evolution*, 2009. **9**(6): pp. 1381–1391.

[8] Xia, R., et al., The *aro*Q-encoded monofunctional chorismate mutase (CM-F) protein is a periplasmic enzyme in *Erwinia herbicola*. *Journal of Bacteriology*, 1993. **175**: pp. 4729–4737.

[9] Burke, S.A., et al., Detection of molecular diversity in *Bacillus atrophaeus* by amplified fragment length polymorphism analysis. *Applied and Environmental Microbiology*, 2004. **70**(5): pp. 2786–2790.

[10] Gibbons, H.S., et al., Genomic signatures of strain selection and enhancement in *Bacillus atrophaeus* var. *globigii*, a historical biowarfare simulant. *PLoS One*, 2011. **6**(3).

[11] Fritze, D. and R. Pukall, Reclassification of bioindicator strains *Bacillus subtilis* DSM 675 and *Bacillus subtilis* DSM 2277 as *Bacillus atrophaeus*. *International Journal of Systematic and Evolutionary Microbiology*, 2001. **51**(2001): pp. 35–37.

[12] Julian, T.R., et al., Comparison of surface sampling methods for virus recovery from fomites. *Applied and Environmental Microbiology*, 2011. **77**(19): pp. 6918–6925.

[13] Lore, M.B., et al., Performance of conventional and antimicrobial-treated filtering facepiece respirators challenged with biological aerosols. *Journal of Occupational and Environmental Hygiene*, 2012. **9**(2): pp. 69–80.

[14] May, T., K. Tsuruta, and S. Okabe, Exposure of conjugative plasmid carrying *Escherichia coli* biofilms to male-specific bacteriophages. *Isme Journal*, 2011. **5**(4): pp. 771–775.

[15] O'Connell, K.P.O., et al., Real-time fluorogenic reverse transcription-PCR assays for detection of bacteriophage MS2. *Applied and Environmental Microbiology*, 2006. **72**(1): pp. 478–483.

[16] Miyake, T., et al., Grouping of RNA phages based on the template specificity of their replicases. *Proceedings of the National Academy of Science*, 1971. **68**(9): p. 2022–2024.

[17] EPA, *Escherichia coli* K-12 derivatives final risk assessment, U.E.P. Agency, Editor 1997, Biotechnology Program.

[18] Feng, P., S.D. Weagant, and M.A. Grant. Bacteriological analytical manual: chapter 4, enumeration of *Escherichia coli* and the coliform bacteria. 2002 [cited 2011 19 January 2011]; Available from: http://www.fda.gov/Food/ScienceResearch/LaboratoryMethods/BacteriologicalAnalyticalManualBAM/ucm064948.htm.

[19] Kuhnert, P. and J. Frey. Tools for safety assessment: Identification and monitoring of *Escherichia coli* K-12 safety strains. 2010 [cited 2010 2 June 2010]; Available from: http://www.bats.ch/bats/publikationen/1996-1_e.coli/96-1_e-coli_k12.php.

[20] Bader, D.E., et al., Evaluation of AG43 PCR primers for *Erwinia herbicola*, bacteriophage MS2, ovalbumin and *Bacillus subtilis* var. *niger* (BG) using lightcycler based fluorescent assays (LRTR-AG43-1-2000), 2000, The Technical Cooperation Program.

[21] ATCC. Product description. 14 July 2011; Available from: http://www.atcc.org/ATCCAdvancedCatalogSearch/ProductDetails/tabid/452/Default.aspx?ATCCNum=15597&Template=bacteria.

[22] Floyd, R.V., et al., Modulation of ureteric Ca signaling and contractility in humans and rats by uropathogenic *E. coli*. *American Journal of Physiology-Renal Physiology*, 2010. **298**(4): pp. F900–F908.

[23] Kuhnert, P., J. Nicolet, and J. Frey, Rapid and accurate identification of *Escherichia coli* K-12 strains. *Applied and Environmental Microbiology*, 1995. **61**(11): pp. 4125–4139.

[24] Madgwick, P.J. and S.G. Waley, β-Lactamase I from *Bacillus cereus*. Structure and site-directed mutagenesis. *Biochem J*, 1987. **248**(3): pp. 657–62.

[25] Yanisch-Perron, C., J. Vieira, and J. Messing, Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, 1985. **33**(1): pp. 103–19.

[26] Buttner, M.P., et al., Evaluation of two surface sampling methods for detection of *Erwinia herbicola* on a variety of materials by culture and quantitative PCR. *Applied and Environmental Microbiology*, 2007. **73**(11): pp. 3505–3510.

[27] Buttner, M.P., et al., Determination of the efficacy of two building decontamination strategies by surface sampling with culture and quantitative PCR analysis. *Applied and Environmental Microbiology*, 2004. **70**(8): pp. 4740–4747.

[28] Smits, T.H.M., et al., Genomic and phenotypic characterization of a nonpigmented variant of *Pantoea vagans* biocontrol strain C9-1 lacking the 530-kb megaplasmid pPag3 [ABSTRACT ONLY]. *FEMS Microbiology Letters*, 2010. **308**(1): pp. 48–54.

Annex A MS2 assay 4 vs MS2 assay 5

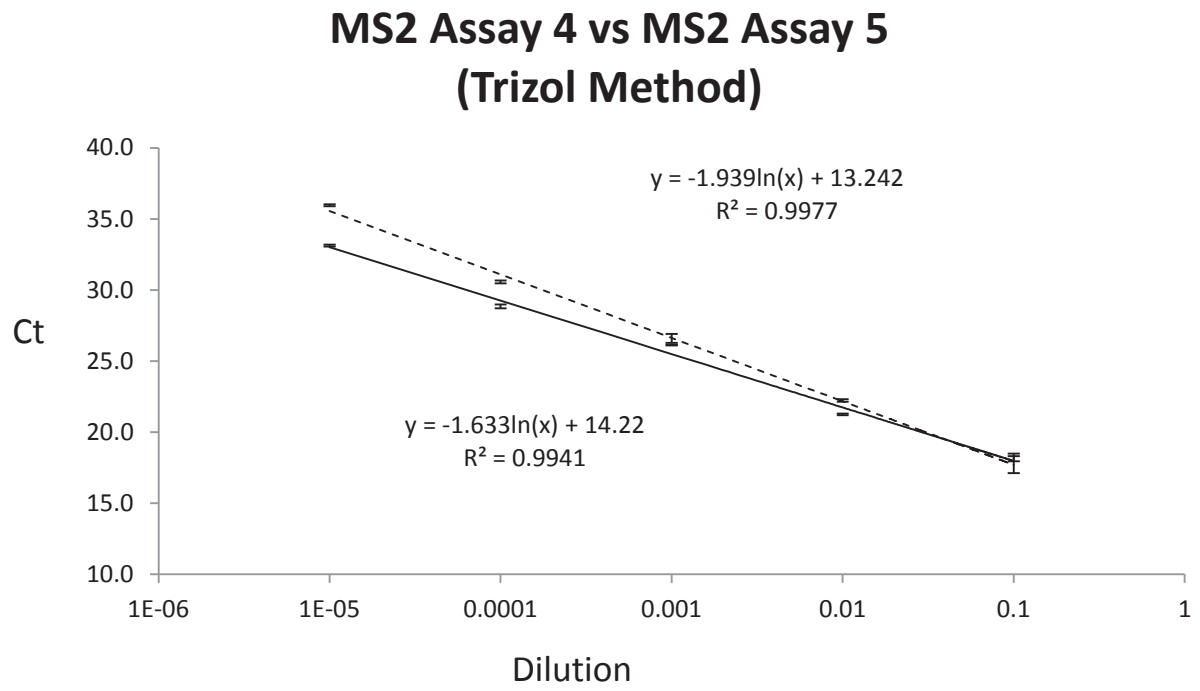

Figure A.1: *PCR standard curve for MS2 assay 4 (dashed line) versus MS2 assay 5 (solid line) using template isolated from MS2 phage using the TRIzol® method.*

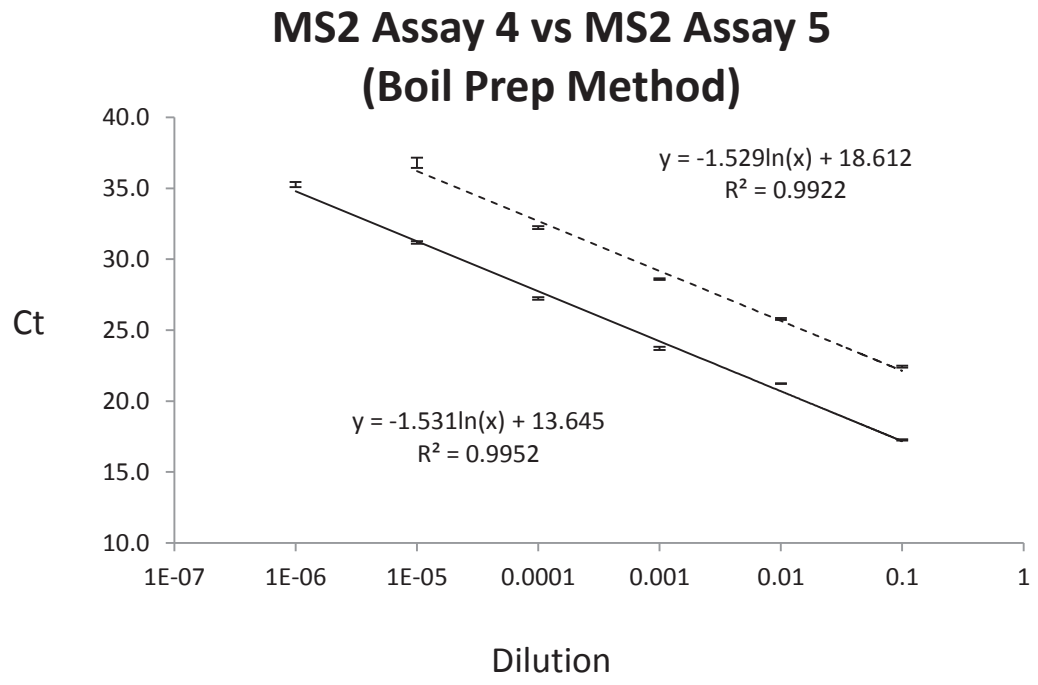

Figure A.2: *PCR standard curve for MS2 assay 4 (dashed line) versus MS2 assay 5 (solid line) using template isolated from MS2 phage using the Boil Prep method.*

Annex B Standard curves for the simulant assays

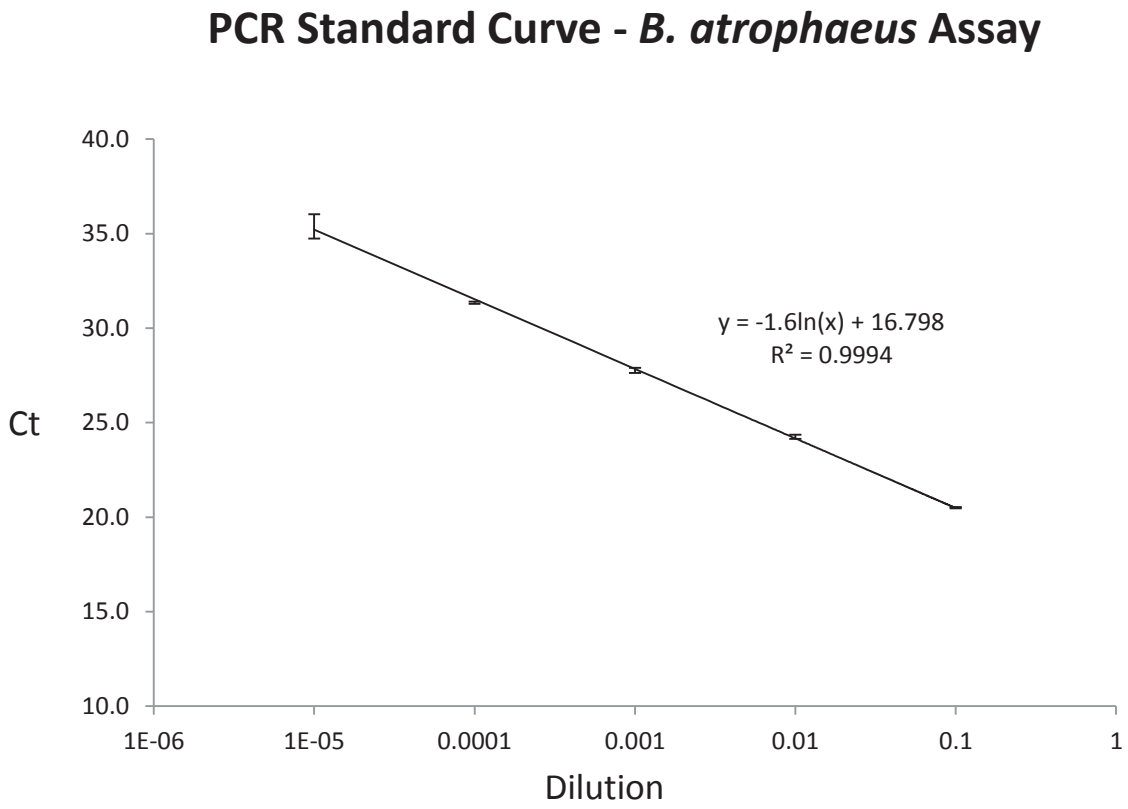

Figure B.1: *PCR standard curve for Bacillus atrophaeus assay using Bacillus atrophaeus positive control template prepared using the Boil Prep method.*

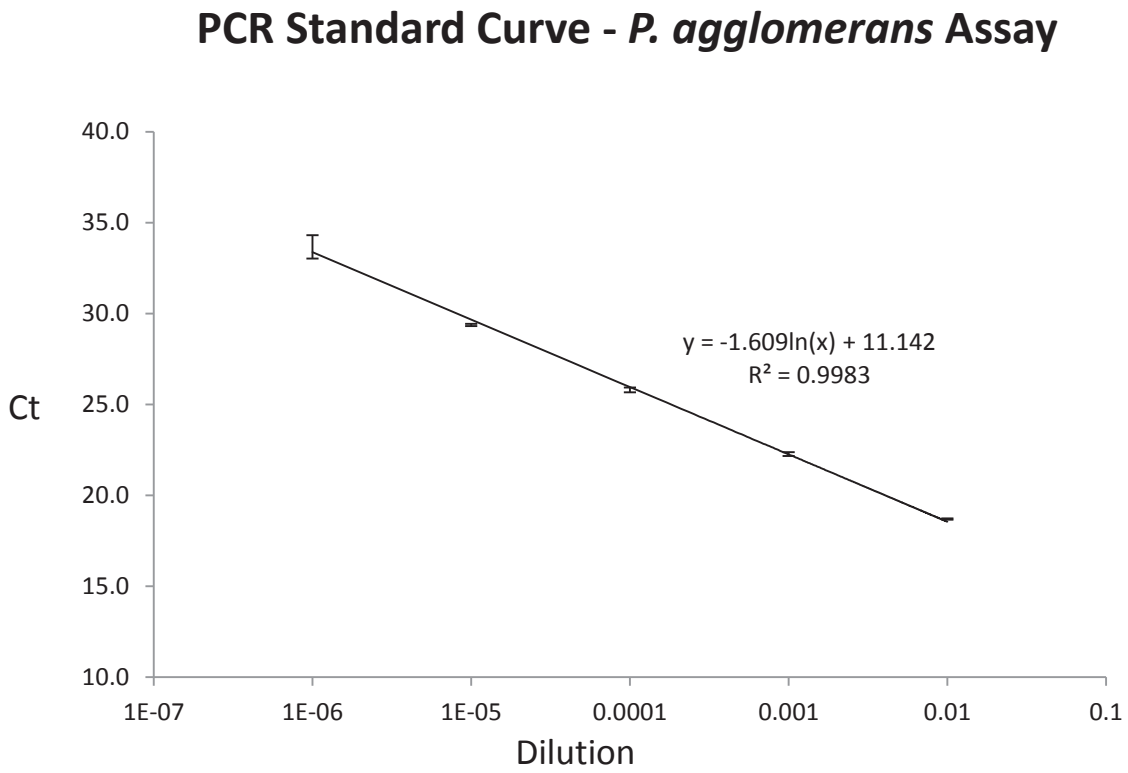

Figure B.2: *PCR standard curve for the Pantoea agglomerans assay using Pantoea agglomerans positive control template prepared using the Boil Prep method.*

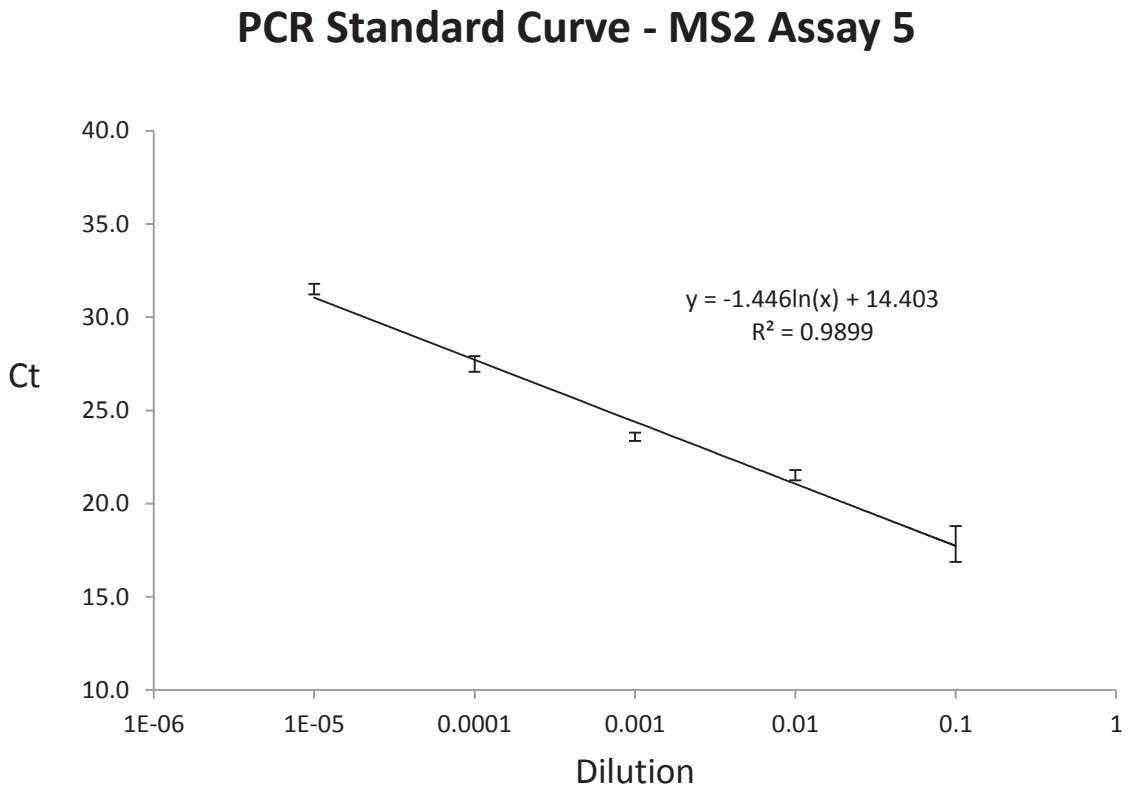

Figure B.3: *PCR standard curve for MS2 assay 5 using MS2 positive control template prepared using the Boil Prep method.*

Annex C Comparison of partial 16S rRNA sequences for *Pantoea agglomerans* E325 and *Pantoea agglomerans* 33243 by Dr Chris Wend

Comparison of partial 16S rRNA sequences for *Pantoea agglomerans* E325 and 33243 04 JUN 2009

Chris Wend, PhD, PE (MT) Director
Bioscience Division
Northwest Agricultural Products Pasco, WA 99301
(509) 851-8667
chrisw@nap-chem.com

In an attempt to determine how closely related E325 and ATCC33243 might be, 16s rRNA partial sequences were analyzed using blastn. The two sequences matched well against each other with an expected value of 0. Further comparison of the NCBI databases for the two organims revealed two 100 organism lists of which 59 were found to be common to both organisms. The 16S rRNA results indicate that these organisms are closely related to similar organisms and one might conjecture that this might indicate a reasonable similarity between E325 and ATCC 33243.

It should be noted that the E325 sequence has a gap in the middle and this results in lower scores.

Respectfully,

Chris Wend

C.1 Blastn comparison of *Pantoea agglomerans* E325 and 33243

Query ID: lcl|60637
Description:gi|e325|gb|U325|E325 *Erwinia herbicola* 16S ribosomal RNA gene, partial sequence
Molecule type: nucleic acid
Query Length:1225
Subject ID: 60639
Description:gi|2570280|gb|U80202.1|EHU80202 *Erwinia herbicola* 16S ribosomal RNA gene, partial seq
Molecule type: nucleic acid
Subject Length: 1454
Program: BLASTN 2.2.20+

Reference

Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schff Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), ”Gapped BLAST and PSI-BLAST: a new generation of protein database search programs”, Nucleic Acids Res. 25:3389-3402.

Search Parameters

Program: blastn
Word size: 11
Expect value: 10
Hitlist size: 100
Match/Mismatch scores: 2,-3
Gapcosts: 5, 2
Low Complexity Filter:Yes
Filter string: L; m;
Genetic Code: 1
Karlin-Altschul statistics (Ungapped; Gapped):

Lambda (0.633731; 0.625)
K (0.408146; 0.41)
H (0.912438; 0.78)

Results Statistics:

Effective search space: 1741190

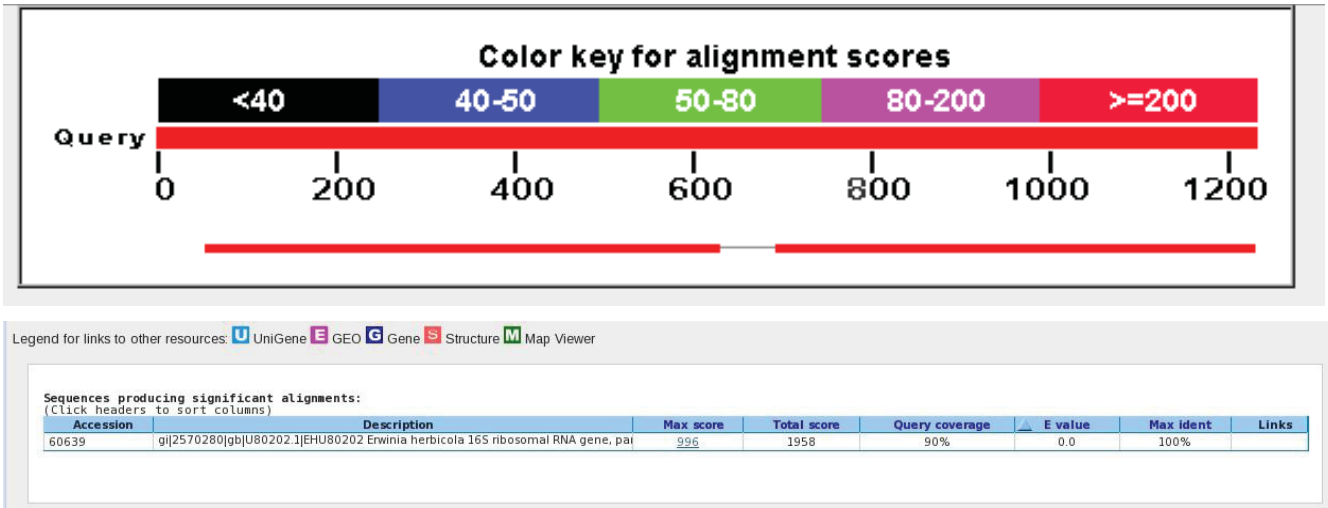

Figure C.1: *Graphic summary of blastn analysis of 16s rRNA partial sequences for Pantoea agglomerans E325 and ATCC 33243.*

Score = 996 bits (1104), Expect = 0.0, Identities = 564/572 (98%), Gaps = 0/572 (0%), Strand=Plus/Minus

C.2 NCBI database search but restricted to *Pantoea agglomerans* database for E325

Database: All GenBank+EMBL+DDBJ+PDB sequences (but no EST, STS, GSS, environmental samples or phase 0, 1 or 2 HTGS sequences)
Query= gi|e325|gb|U325|E325 Erwinia herbicola 16S ribosomal RNA gene, partial sequence
Length=1225
Posted date: Jun 3, 2009 5:42 PM
Number of letters in database: 847,927
Number of sequences in database: 634
Lambda, K, H: 0.634, 0.408, 0.912
Gapped Lambda, K, H: 0.625, 0.410, 0.780
Matrix: blastn
matrix: 2, -3
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 634
Number of Hits to DB: 661
Number of extensions: 24
Number of successful extensions: 14
Number of sequences better than 10: 5
Number of HSP's better than 10 without gapping: 0
Number of HSP's gapped: 8
Number of HSP's successfully gapped: 8
Length of query: 1225
Length of database: 847927
Length adjustment: 23
Effective length of query: 1202
Effective length of database: 833345
Effective search space: 1001680690
Effective search space used: 1001680690

Table C.1: *NCBI sequences producing significant alignments to E325 but restricted to Pantoea agglomerans database.*

Sequences producing significant alignments		Score (Bits)	E Value
gb|FJ756355.1|	Pantoea agglomerans strain 102 16S ribosomal R...	1032	0.0
gb|FJ756348.1|	Pantoea agglomerans strain NZ 16S ribosomal RN...	1032	0.0
gb|FJ611851.1|	Pantoea agglomerans strain P4SAH 16S ribosomal...	1032	0.0
gb|FJ611843.1|	Pantoea agglomerans strain EM21cb 16S ribosoma...	1032	0.0
gb|FJ611842.1|	Pantoea agglomerans strain Eh325 16S ribosomal...	1032	0.0
gb|FJ611841.1|	Pantoea agglomerans strain EPS 125 16S ribosom...	1032	0.0
gb|FJ611840.1|	Pantoea agglomerans strain ACW 55899 16S ribos...	1032	0.0
gb|FJ611839.1|	Pantoea agglomerans strain LMG 1286 16S riboso...	1032	0.0
gb|FJ611835.1|	Pantoea agglomerans strain LMG 2595 16S riboso...	1032	0.0

Sequences producing significant alignments		Score (Bits)	E Value
gb|FJ611834.1|	Pantoea agglomerans strain CPA-2 16S ribosomal...	1032	0.0
gb|FJ611833.1|	Pantoea agglomerans strain VA21971 16S ribosom...	1032	0.0
gb|FJ611822.1|	Pantoea agglomerans strain P5WAM 16S ribosomal...	1032	0.0
gb|FJ611821.1|	Pantoea agglomerans pv. gypsophilae strain ATC...	1032	0.0
gb|EU272867.1|	Pantoea agglomerans strain MK26 16S ribosomal ...	1032	0.0
gb|FJ357813.1|	Pantoea agglomerans strain BBPE8554 16S riboso...	1032	0.0
gb|FJ357809.1|	Pantoea agglomerans strain BBPE0103202 16S rib...	1032	0.0
gb|AF130952.2|	Enterobacter agglomerans strain A92 16S riboso...	1032	0.0
gb|AF130934.2|	Enterobacter agglomerans strain A65 16S riboso...	1032	0.0
gb|AF130927.2|	Enterobacter agglomerans strain A57 16S riboso...	1032	0.0
gb|EU598802.1|	Pantoea agglomerans 16S ribosomal RNA gene, pa...	1032	0.0
gb|EU047555.1|	Pantoea agglomerans strain TMPSB-P1 16S riboso...	1032	0.0
gb|DQ307453.1|	Pantoea agglomerans strain MM2 16S ribosomal R...	1032	0.0
gb|DQ307452.1|	Pantoea agglomerans strain FL1 16S ribosomal R...	1032	0.0
dbj|AB004757.2|	Pantoea agglomerans gene for 16S ribosomal RN...	1032	0.0
emb|AJ583835.1|	Pantoea agglomerans partial 16S rRNA gene, st...	1032	0.0
emb|AJ583011.1|	Pantoea agglomerans partial 16S rRNA gene, st...	1032	0.0
gb|AF130946.1|	Enterobacter agglomerans strain A81 16S riboso...	1032	0.0
gb|AF130944.1|	Enterobacter agglomerans strain A79 16S riboso...	1032	0.0
gb|AF130943.1|	Enterobacter agglomerans strain A77 16S riboso...	1032	0.0
gb|AF130895.1|	Pantoea agglomerans strain new*16 16S ribosoma...	1032	0.0
emb|AJ233423.1|	Pantoea agglomerans 16S rRNA gene (strain DSM...	1032	0.0
emb|Z96082.1|	Pantoea agglomerans LMG 2565 16S ribosomal RNA	1032	0.0
gb|FJ611831.1|	Pantoea agglomerans strain P9QLB 16S ribosomal...	1029	0.0
gb|FJ357814.1|	Pantoea agglomerans strain CUETM8553 16S ribos...	1029	0.0
emb|AJ251466.1|	Pantoea agglomerans partial16S rRNA gene, st...	1029	0.0
gb|AF130924.1|	Enterobacter agglomerans strain A54 16S riboso...	1029	0.0
gb|FJ611838.1|	Pantoea agglomerans strain EM22cb 16S ribosoma...	1027	0.0
gb|FJ611832.1|	Pantoea agglomerans strain CIP 82.100 16S ribo...	1027	0.0
gb|FJ611830.1|	Pantoea agglomerans strain P11QLN 16S ribosoma...	1027	0.0
gb|FJ357811.1|	Pantoea agglomerans strain BBPE277471 16S ribo...	1027	0.0
gb|AF373197.1|	Pantoea agglomerans strain GSPB 450 16S riboso...	1027	0.0
gb|AF373196.1|	Pantoea agglomerans strain LMG 2565 16S riboso...	1027	0.0
emb|AM184091.1|	Pantoea agglomerans partial16S rRNA gene, st...	1027	0.0
gb|FJ611837.1|	Pantoea agglomerans strain LMG 2941 16S riboso...	1023	0.0
gb|FJ611836.1|	Pantoea agglomerans strain P10c 16S ribosomal ...	1023	0.0
gb|EU834421.1|	Pantoea agglomerans strain R2-569 16S ribosoma...	1023	0.0
gb|AY741162.1|	Pantoea agglomerans strain S33 16S R...ribosomal	1023	0.0

Sequences producing significant alignments		Score (Bits)	E Value
gb|DQ365572.1|	Pantoea agglomerans strain JA07 16S ...ribosomal	1023	0.0
gb|AF130951.1|	Enterobacter agglomerans strain A91 16S riboso...	1023	0.0
gb|AF130933.1|	Enterobacter agglomerans strain A63 16S riboso...	1023	0.0
gb|U80183.1|EMU80183	Erwinia milletiae 16S ribosomal RNA gene...	1023	0.0
gb|DQ065752.1|	Pantoea agglomerans strain I10 16S R...ribosomal	1021	0.0
gb|AF130930.1|	Enterobacter agglomerans strain A60 16S riboso...	1021	0.0
gb|AF130922.1|	Enterobacter agglomerans strain A51 16S riboso...	1021	0.0
gb|FJ357812.1|	Pantoea agglomerans strain BBPE8284 16S riboso...	1020	0.0
gb|AY530797.1|	Pantoea agglomerans strain SUH 2 16S ribosomal...	1020	0.0
gb|AY941841.1|	Pantoea agglomerans strain XW131 16S ribosomal...	1020	0.0
gb|AF130949.1|	Enterobacter agglomerans strain A87 16S riboso...	1020	0.0
gb|FJ611829.1|	Pantoea agglomerans pv. gypsophilae strain CFB...	1018	0.0
gb|DQ307454.1|	Pantoea agglomerans strain SC1 16S ribosomal R...	1018	0.0
gb|AY941838.1|	Pantoea agglomerans strain XW123 16S ribosomal...	1018	0.0
gb|AF157694.1|AF157694	Pantoea agglomerans16S ribosomal RNA ...	1018	0.0
gb|AF130900.1|	Enterobacter agglomerans strain A22 16S riboso...	1018	0.0
gb|EF050806.1|	Pantoea agglomerans strain PGHLT4 16S ribosoma...	1016	0.0
gb|AF130935.1|	Enterobacter agglomerans strain A67 16S riboso...	1016	0.0
gb|AF130932.1|	Enterobacter agglomerans strain A62 16S riboso...	1016	0.0
gb|AF130923.1|	Enterobacter agglomerans strain A53 16S riboso...	1016	0.0
gb|AF130901.1|	Enterobacter agglomerans strain A23 16S riboso...	1016	0.0
gb|FJ756346.1|	Pantoea agglomerans strain LMG 2734 16S riboso...	1014	0.0
gb|FJ357810.1|	Pantoea agglomerans strain BBPE014230 16S ribo...	1014	0.0
gb|EU849107.1|	Pantoea agglomerans strain U2-21 16S ribosomal...	1014	0.0
gb|EU130700.1|	Pantoea agglomerans strain PGHL23-15 16S ribos...	1014	0.0
gb|EF178448.1|	Pantoea agglomerans strain2Re40 16S ribosomal...	1014	0.0
gb|FJ611828.1|	Pantoea agglomerans strain Eh460 16S ribosomal...	1012	0.0
emb|AM293681.1|	Pantoea agglomerans partial16S rRNA gene, st...	1012	0.0
gb|FJ611827.1|	Pantoea agglomerans strain Eh1087 16S ribosoma...	1011	0.0
gb|FJ611819.1|	Pantoea agglomerans strain P2SAA 16S ribosomal...	1011	0.0
gb|FJ611818.1|	Pantoea agglomerans strain P1SAA 16S ribosomal...	1011	0.0
gb|AF130906.2|	Enterobacter agglomerans strain A28 16S riboso...	1011	0.0
gb|AF130937.1|	Enterobacter agglomerans strain A69 16S riboso...	1011	0.0
gb|AF130897.1|	Enterobacter agglomerans strain A18 16S riboso...	1011	0.0
gb|FJ611826.1|	Pantoea agglomerans strain Eh239 16S ribosomal...	1009	0.0
gb|FJ611823.1|	Pantoea agglomerans strain P7NSW 16S ribosomal...	1009	0.0
gb|FJ611820.1|	Pantoea agglomerans strain P3SAA 16S ribosomal...	1009	0.0
gb|AF130939.1|	Enterobacter agglomerans strain A71 16S riboso...	1009	0.0

Sequences producing significant alignments		Score (Bits)	E Value
gb|AF130904.1|	Enterobacter agglomerans strain A26 16S riboso...	1009	0.0
gb|AF130898.1|	Enterobacter agglomerans strain A19 16S riboso...	1009	0.0
emb|FN263076.1|	Pantoea agglomerans partial 16S rRNA gene, st...	1007	0.0
gb|FJ611825.1|	Pantoea agglomerans strain Eh454 16S ribosomal...	1007	0.0
gb|AF130925.1|	Enterobacter agglomerans strain A55 16S riboso...	1007	0.0
emb|Z96083.1|	Pantoea agglomerans LMG 2660 16S ribosomal RNA	1007	0.0
gb|AY395010.1|	Pantoea agglomerans 16S ribosomal RNA gene, pa...	1005	0.0
gb|FJ611844.1|	Pantoea agglomerans strain P6WAL 16S ribosomal...	1003	0.0
gb|FJ611824.1|	Pantoea agglomerans strain ATCC 27987 16S ribo...	1003	0.0
gb|AF130953.1|	Pantoea agglomerans strain ATCC 27155 16S ribo...	1003	0.0
gb|EU849108.1|	Pantoea agglomerans strain U2-22 16S ribosomal...	1000	0.0
gb|EF050810.1|	Pantoea agglomerans strain PGHL6 16S ribosomal...	1000	0.0
gb|AF130938.1|	Enterobacter agglomerans strain A70 16S riboso...	1000	0.0
gb|AF130931.1|	Enterobacter agglomerans strain A61 16S riboso...	998	0.0
gb|AF130886.2|	Enterobacter agglomerans strain A126 16S ribos...	993	0.0

C.3 NCBI database search but restricted to *Pantoea agglomerans* database for ATCC 33243

Database: All GenBank+EMBL+DDBJ+PDB sequences (but no EST, STS, GSS, environmental samples or phase 0, 1 or 2 HTGS sequences)
Query= gi|2570280|gb|U80202.1|EHU80202 Erwinia herbicola 16S ribosomal RNA gene, partial sequence
Length=1454
Posted date: Jun 3, 2009 5:42 PM
Number of letters in database: 847,927
Number of sequences in database: 634
Lambda, K, H: 0.634, 0.408, 0.912
Gapped Lambda, K, H: 0.625, 0.410, 0.780
Matrix: blastn
matrix: 2, -3
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 634
Number of Hits to DB: 930
Number of extensions: 25
Number of successful extensions: 20
Number of sequences better than 10: 5
Number of HSP's better than 10 without gapping: 0
Number of HSP's gapped: 5
Number of HSP's successfully gapped: 5
Length of query: 1454
Length of database: 847927
Length adjustment: 23
Effective length of query: 1431
Effective length of database: 833345
Effective search space: 1192516695
Effective search space used: 1192516695

Table C.2: *NCBI sequences producing significant alignments to ATTC 33243 but restricted to Pantoea agglomerans database.*

Sequences producing significant alignments		Score (Bits)	E Value
gb|U80202.1|EHU80202	Erwinia herbicola 16S ribosomal RNA gene...	2623	0.0
gb|FJ756354.1|	Pantoea agglomerans strain 48b/90 16S ribosoma...	2603	0.0
gb|AY395010.1|	Pantoea agglomerans 16S ribosomal RNA gene, pa...	2594	0.0
gb|EU849108.1|	Pantoea agglomerans strain U2-22 16S ribosomal...	2585	0.0
gb|FJ357815.1|	Pantoea agglomerans strain EL107 16S ribosomal...	2583	0.0
gb|AF130916.1|	Pantoea agglomerans strain new*45con 16S ribos...	2581	0.0
gb|AY849936.1|	Pantoea agglomerans strain BJ-Tobacco 16S ribo...	2576	0.0
gb|AF130884.1|	Enterobacter agglomerans strain A123 16S ribos...	2574	0.0
gb|FJ756347.1|	Pantoea agglomerans strain HY 5080 16S ribosom...	2571	0.0

Sequences producing significant alignments		Score (Bits)	E Value
gb|AF130945.2|	Enterobacter agglomerans strain A80 16S riboso...	2567	0.0
gb|EU598802.1|	Pantoea agglomerans 16S ribosomal RNA gene, pa...	2567	0.0
gb|AY924375.1|	Pantoea agglomerans strain Sc-2 16S ribosomal ...	2567	0.0
gb|AY924374.1|	Pantoea agglomerans strain Sc-1 16S ribosomal ...	2567	0.0
gb|AF130887.1|	Enterobacter agglomerans strain A20 16S riboso...	2567	0.0
gb|AF130928.2|	Enterobacter agglomerans strain A58 16S riboso...	2565	0.0
gb|AF130948.2|	Enterobacter agglomerans strain A84 16S riboso...	2563	0.0
gb|AF130941.2|	Enterobacter agglomerans strain A73 16S riboso...	2563	0.0
gb|FJ756356.1|	Pantoea agglomerans strain EhY112-9/86 16S rib...	2562	0.0
gb|FJ357810.1|	Pantoea agglomerans strain BBPE014230 16S ribo...	2562	0.0
gb|AF130947.2|	Enterobacter agglomerans strain A83 16S riboso...	2562	0.0
gb|AF130934.2|	Enterobacter agglomerans strain A65 16S riboso...	2562	0.0
gb|AY924376.1|	Pantoea agglomerans strain Sc-4 16S ribosomal ...	2562	0.0
gb|AF130938.1|	Enterobacter agglomerans strain A70 16S riboso...	2562	0.0
gb|AF130929.1|	Enterobacter agglomerans strain A59 16S riboso...	2560	0.0
gb|AF130907.1|	Enterobacter agglomerans strain A29 16S riboso...	2560	0.0
gb|FJ357813.1|	Pantoea agglomerans strain BBPE8554 16S riboso...	2556	0.0
gb|FJ357809.1|	Pantoea agglomerans strain BBPE0103202 16S rib...	2556	0.0
emb|AM184097.1|	Pantoea agglomerans partial 16S rRNA gene, st...	2556	0.0
gb|AF130912.1|	Enterobacter agglomerans strain A40 16S riboso...	2556	0.0
gb|EU047555.1|	Pantoea agglomerans strain TMPSB-P1 16S riboso...	2554	0.0
emb|AM184091.1|	Pantoea agglomerans partial 16S rRNA gene, st...	2554	0.0
gb|FJ357811.1|	Pantoea agglomerans strain BBPE277471 16S ribo...	2553	0.0
gb|AF130927.2|	Enterobacter agglomerans strain A57 16S riboso...	2553	0.0
gb|AF130905.1|	Enterobacter agglomerans strain A27 16S riboso...	2551	0.0
gb|U80183.1|EMU80183	Erwinia milletiae 16S ribosomal RNA gene...	2547	0.0
emb|AM234150.1|	Pantoea agglomerans partial 16S rRNA gene, st...	2545	0.0
gb|AF130949.1|	Enterobacter agglomerans strain A87 16S riboso...	2545	0.0
emb|AJ233423.1|	Pantoea agglomerans 16S rRNA gene (strain DSM...	2545	0.0
gb|AF130944.1|	Enterobacter agglomerans strain A79 16S riboso...	2542	0.0
gb|FJ756346.1|	Pantoea agglomerans strain LMG 2734 16S riboso...	2540	0.0
gb|AF373196.1|	Pantoea agglomerans strain LMG 2565 16S riboso...	2540	0.0
gb|AY941841.1|	Pantoea agglomerans strain XW131 16S ribosomal...	2540	0.0
emb|AJ251466.1|	Pantoea agglomerans partial 16S rRNA gene, st...	2540	0.0
gb|AF130930.1|	Enterobacter agglomerans strain A60 16S riboso...	2540	0.0
gb|AF130918.1|	Pantoea agglomerans strain new*47con 16S ribos...	2540	0.0
gb|AF130911.1|	Enterobacter agglomerans strain A38 16S riboso...	2536	0.0
gb|AF130895.1|	Pantoea agglomerans strain new*16 16S ribosoma...	2536	0.0
gb|AF130937.1|	Enterobacter agglomerans strain A69 16S riboso...	2535	0.0
gb|AF130924.1|	Enterobacter agglomerans strain A54 16S riboso...	2533	0.0

Sequences producing significant alignments		Score (Bits)	E Value
gb|DQ365572.1|	Pantoea agglomerans strain JA07 16S ribosomal ...	2531	0.0
gb|AF130940.1|	Enterobacter agglomerans strain A72 16S riboso...	2529	0.0
gb|AF130917.1|	Enterobacter agglomerans strain A46 16S riboso...	2529	0.0
gb|AY941838.1|	Pantoea agglomerans strain XW123 16S ribosomal...	2527	0.0
gb|AF130946.1|	Enterobacter agglomerans strain A81 16S riboso...	2527	0.0
emb|Z96083.1|	Pantoea agglomerans LMG 2660 16S ribosomal RNA	2527	0.0
gb|AF130952.2|	Enterobacter agglomerans strain A92 16S riboso...	2524	0.0
gb|AF130901.1|	Enterobacter agglomerans strain A23 16S riboso...	2524	0.0
emb|FN263076.1|	Pantoea agglomerans partial 16S rRNA gene, st...	2520	0.0
dbj|AB004757.2|	Pantoea agglomerans gene for 16S ribosomal RN...	2520	0.0
gb|AF130936.1|	Enterobacter agglomerans strain A68 16S riboso...	2520	0.0
gb|AF130933.1|	Enterobacter agglomerans strain A63 16S riboso...	2520	0.0
gb|DQ307454.1|	Pantoea agglomerans strain SC1 16S ribosomal R...	2518	0.0
gb|AF130925.1|	Enterobacter agglomerans strain A55 16S riboso...	2518	0.0
gb|EU849107.1|	Pantoea agglomerans strain U2-21 16S ribosomal...	2513	0.0
gb|AF130951.1|	Enterobacter agglomerans strain A91 16S riboso...	2513	0.0
gb|AF130931.1|	Enterobacter agglomerans strain A61 16S riboso...	2511	0.0
gb|FJ756348.1|	Pantoea agglomerans strain NZ 16S ribosomal RN...	2509	0.0
gb|AF130939.1|	Enterobacter agglomerans strain A71 16S riboso...	2509	0.0
gb|AF130935.1|	Enterobacter agglomerans strain A67 16S riboso...	2507	0.0
gb|AF130899.1|	Enterobacter agglomerans strain A21 16S riboso...	2507	0.0
gb|AF130942.1|	Enterobacter agglomerans strain A74 16S riboso...	2506	0.0
gb|AY741162.1|	Pantoea agglomerans strain S33 16S ribosomal R...	2504	0.0
gb|AY530797.1|	Pantoea agglomerans strain SUH 2 16S ribosomal...	2504	0.0
gb|DQ307452.1|	Pantoea agglomerans strain FL1 16S ribosomal R...	2500	0.0
gb|AF130913.1|	Enterobacter agglomerans strain A41 16S riboso...	2498	0.0
gb|FJ357816.1|	Pantoea agglomerans strain EM102 16S ribosomal...	2497	0.0
gb|AF130922.1|	Enterobacter agglomerans strain A51 16S riboso...	2495	0.0
gb|AF130923.1|	Enterobacter agglomerans strain A53 16S riboso...	2493	0.0
gb|EU272867.1|	Pantoea agglomerans strain MK26 16S ribosomal ...	2491	0.0
gb|FJ357834.1|	Pantoea agglomerans strain SB545 16S ribosomal...	2491	0.0
gb|DQ307453.1|	Pantoea agglomerans strain MM2 16S ribosomal R...	2488	0.0
emb|Z96082.1|	Pantoea agglomerans LMG 2565 16S ribosomal RNA	2488	0.0
gb|AF130961.1|	Enterobacter agglomerans strain A9 16S ribosom...	2486	0.0
gb|AF130932.1|	Enterobacter agglomerans strain A62 16S riboso...	2486	0.0
gb|FJ357812.1|	Pantoea agglomerans strain BBPE8284 16S riboso...	2484	0.0
gb|AF130943.1|	Enterobacter agglomerans strain A77 16S riboso...	2484	0.0
gb|AF130926.1|	Enterobacter agglomerans strain A56 16S riboso...	2477	0.0
gb|AY092079.1|	Pantoea agglomerans strain 732 16S ribosomal R...	2475	0.0
gb|AF373197.1|	Pantoea agglomerans strain GSPB 450 16S riboso...	2475	0.0

Sequences producing significant alignments		Score (Bits)	E Value
gb|FJ357814.1|	Pantoea agglomerans strain CUETM8553 16S ribos...	2471	0.0
emb|FM202484.1|	Pantoea agglomerans partial 16S rRNA gene, st...	2470	0.0
emb|AJ583011.1|	Pantoea agglomerans partial 16S rRNA gene, st...	2455	0.0
gb|FJ611809.1|	Pantoea agglomerans strain LMG2558 16S ribosom...	2452	0.0
gb|AF130898.1|	Enterobacter agglomerans strain A19 16S riboso...	2452	0.0
gb|FJ611821.1|	Pantoea agglomerans pv. gypsophilae strain ATC...	2450	0.0
emb|FM202486.1|	Pantoea agglomerans partial 16S rRNA gene, st...	2448	0.0
gb|AF157694.1|AF157694	Pantoea agglomerans 16S ribosomal RNA ...	2446	0.0
emb|AJ583835.1|	Pantoea agglomerans partial 16S rRNA gene, st...	2444	0.0
gb|FJ611802.1|	Pantoea agglomerans strain LMG 2557 16S riboso...	2437	0.0
gb|FJ788420.1|	Pantoea agglomerans strain CLJ1 16S ribosomal ...	2434	0.0

Table C.3: *List of 59 organisms common to Pantoea agglomerans E325 and ATCC 33243 with their corresponding bit scores.*

Sequences producing significant alignments		E325 Score (Bits)	ATCC 33243 Score (Bits)	E Value
FJ756348.1	Pantoea agglomerans strain NZ 16S ribosomal RN...	1032	2509	0.0
FJ611821.1	Pantoea agglomerans pv. gypsophilae strain ATC...	1032	2450	0.0
EU272867.1	Pantoea agglomerans strain MK26 16S ribosomal ...	1032	2491	0.0
FJ357813.1	Pantoea agglomerans strain BBPE8554 16S riboso...	1032	2556	0.0
FJ357809.1	Pantoea agglomerans strain BBPE0103202 16S rib...	1032	2556	0.0
AF130952.2	Enterobacter agglomerans strain A92 16S riboso...	1032	2524	0.0
AF130934.2	Enterobacter agglomerans strain A65 16S riboso...	1032	2562	0.0
AF130927.2	Enterobacter agglomerans strain A57 16S riboso...	1032	2553	0.0
EU598802.1	Pantoea agglomerans 16S ribosomal RNA gene, pa...	1032	2567	0.0
EU047555.1	Pantoea agglomerans strain TMPSB-P1 16S riboso...	1032	2554	0.0
DQ307453.1	Pantoea agglomerans strain MM2 16S ribosomal R...	1032	2488	0.0
DQ307452.1	Pantoea agglomerans strain FL1 16S ribosomal R...	1032	2500	0.0
AB004757.2	Pantoea agglomerans gene for 16S ribosomal RN...	1032	2520	0.0
AJ583835.1	Pantoea agglomerans partial 16S rRNA gene, st...	1032	2444	0.0
AJ583011.1	Pantoea agglomerans partial 16S rRNA gene, st...	1032	2455	0.0
F130946.1	Enterobacter agglomerans strain A81 16S riboso...	1032	2527	0.0
AF130944.1	Enterobacter agglomerans strain A79 16S riboso...	1032	2542	0.0
AF130943.1	Enterobacter agglomerans strain A77 16S riboso...	1032	2484	0.0
AF130895.1	Pantoea agglomerans strain new*16 16S ribosoma...	1032	2536	0.0
AJ233423.1	Pantoea agglomerans 16S rRNA gene (strain DSM...	1032	2545	0.0
Z96082.1	Pantoea agglomerans LMG 2565 16S ribosomal RNA	1032	2488	0.0
FJ357814.1	Pantoea agglomerans strain CUETM8553 16S ribos...	1029	2471	0.0
AJ251466.1	Pantoea agglomerans partial 16S rRNA gene, st...	1029	2540	0.0
AF130924.1	Enterobacter agglomerans strain A54 16S riboso...	1029	2533	0.0
FJ357811.1	Pantoea agglomerans strain BBPE277471 16S ribo...	1027	2553	0.0
AF373197.1	Pantoea agglomerans strain GSPB 450 16S riboso...	1027	2475	0.0
AF373196.1	Pantoea agglomerans strain LMG 2565 16S riboso...	1027	2540	0.0
AM184091.1	Pantoea agglomerans partial 16S rRNA gene, st...	1027	2554	0.0
AY741162.1	Pantoea agglomerans strain S33 16S ribosomal R...	1023	2504	0.0
DQ365572.1	Pantoea agglomerans strain JA07 16S ribosomal ...	1023	2531	0.0
AF130951.1	Enterobacter agglomerans strain A91 16S riboso...	1023	2513	0.0
AF130933.1	Enterobacter agglomerans strain A63 16S riboso...	1023	2520	0.0
U80183.1 EMU80183	Erwinia milletiae 16S ribosomal RNA gene...	1023	2547	0.0
AF130930.1	Enterobacter agglomerans strain A60 16S riboso...	1021	2540	0.0

Sequences producing significant alignments		E325 Score (Bits)	ATCC 33243 Score (Bits)	E Value
AF130922.1	Enterobacter agglomerans strain A51 16S riboso...	1021	2495	0.0
FJ357812.1	Pantoea agglomerans strain BBPE8284 16S riboso...	1020	2484	0.0
AY530797.1	Pantoea agglomerans strain SUH 2 16S ribosomal...	1020	2504	0.0
AY941841.1	Pantoea agglomerans strain XW131 16S ribosomal...	1020	2545	0.0
AF130949.1	Enterobacter agglomerans strain A87 16S riboso...	1020	2509	0.0
DQ307454.1	Pantoea agglomerans strain SC1 16S ribosomal R...	1018	2518	0.0
AY941838.1	Pantoea agglomerans strain XW123 16S ribosomal...	1018	2527	0.0
AF157694.1 AF157694	Pantoea agglomerans 16S ribosomal RNA ...	1018	2446	0.0
AF130935.1	Enterobacter agglomerans strain A67 16S riboso...	1016	2507	0.0
AF130932.1	Enterobacter agglomerans strain A62 16S riboso...	1016	2486	0.0
AF130923.1	Enterobacter agglomerans strain A53 16S riboso...	1016	2493	0.0
AF130901.1	Enterobacter agglomerans strain A23 16S riboso...	1016	2524	0.0
FJ756346.1	Pantoea agglomerans strain LMG 2734 16S riboso...	1014	2540	0.0
FJ357810.1	Pantoea agglomerans strain BBPE014230 16S ribo...	1014	2562	0.0
EU849107.1	Pantoea agglomerans strain U2-21 16S ribosomal...	1014	2513	0.0
AF130937.1	Enterobacter agglomerans strain A69 16S riboso...	1011	2535	0.0
AF130939.1	Enterobacter agglomerans strain A71 16S riboso...	1009	2509	0.0
AF130898.1	Enterobacter agglomerans strain A19 16S riboso...	1009	2452	0.0
FN263076.1	Pantoea agglomerans partial 16S rRNA gene, st...	1007	2520	0.0
AF130925.1	Enterobacter agglomerans strain A55 16S riboso...	1007	2518	0.0
Z96083.1	Pantoea agglomerans LMG 2660 16S ribosomal RNA	1007	2527	0.0
AY395010.1	Pantoea agglomerans 16S ribosomal RNA gene, pa...	1005	2594	0.0
EU849108.1	Pantoea agglomerans strain U2-22 16S ribosomal...	1000	2585	0.0
AF130938.1	Enterobacter agglomerans strain A70 16S riboso...	1000	2562	0.0
AF130931.1	Enterobacter agglomerans strain A61 16S riboso...	998	2511	0.0

Annex D NCBI genbank sequence data for select *P. agglomerans* strains

D.1 *Pantoea agglomerans* strain Eh325 16S ribosomal RNA gene, partial sequence

GenBank number: FJ611842.1
Downloaded from : http://www.ncbi.nlm.nih.gov/nuccore/FJ611842.1 (27 April 2011)

```
LOCUS       FJ611842     1360 bp    DNA     linear   BCT 24-NOV-2009

DEFINITION  Pantoea agglomerans strain Eh325 16S ribosomene, partial
sequence.
ACCESSION   FJ611842
VERSION     FJ611842.1  GI:223557746
KEYWORDS    .
SOURCE      Pantoea agglomerans
ORGANISM  Pantoea agglomerans Bacteria; Proteobacteria; Gammaproteobacteria;
Enterobacteriales; Enterobacteriaceae; Pantoea.
REFERENCE   1  (bases 1 to 1360)
AUTHORS   Rezzonico,F., Smits,T.H., Montesinos,E., Frey,J.E. and Duffy,B.
TITLE     Genotypic comparison of Pantoea agglomerans plant and clinical
strains
JOURNAL   BMC Microbiol. 9, 204 (2009)
PUBMED   19772624
REMARK    Publication Status: Online-Only
REFERENCE   2  (bases 1 to 1360)
AUTHORS   Rezzonico,F., Smits,T.H.M., Montesinos,E., Frey,J.E. and Duffy,B.
TITLE     Direct Submission
JOURNAL   Submitted (08-JAN-2009) Plant Protection Division, Agroscope
Changins-Waedenswil ACW, Schloss, Postfach 185, Waedenswil 8820, Switzerland
FEATURES             Location/Qualifiers
     source          1..1360
                     /organism="Pantoea agglomerans"
                     /mol_type="genomic DNA"
                     /strain="Eh325"
                     /db_xref="taxon:549"
     rRNA            <1..>1360
                     /product="16S ribosomal RNA"
```

```
   1 ggacgggtga gtaatgtctg gggatctgcc cgatagaggg ggataaccac tggaaacggt
  61 ggctaatacc gcataacgtc gcaagaccaa agaggggggac cttcgggcct ctcactatcg
 121 gatgaaccca gatgggatta gctagtaggc ggggtaatgg cccacctagg cgacgatccc
 181 tagctggtct gagaggatga ccagccacac tggaactgag acacggtcca gactcctacg
 241 ggaggcagca gtggggaata ttgcacaatg ggcgcaagcc tgatgcagcc atgccgcgtg
 301 tatgaagaag gccttcgggt tgtaaagtac tttcagcggg gaggaaggcg aygsggttaa
 361 taaccgcgtc gattgacgtt acccgcagaa gaagcrccgg ctaactccgt gccagcascc
 421 gcggtaatac ggagggtgca agcgttaatc ggaattactg ggcgtaaagc gcacgcaggc
 481 ggtctgttaa gtcagatgtg aaatccccgg gcttaacctg ggaactgcat ttgaaactgg
 541 caggcttgag tcttgtagag gggggtagaa ttccaggtgt agcggtgaaa tgcgtagaga
 601 tctggaggaa taccggtggc gaaggcggcc ccctggacaa agactgacgc tcaggtgcga
 661 aagcgtgggg agcaaacagg attagatacc ctggtagtcc acgccgtaaa cgatgtcgac
 721 ttggaggttg ttcccttgag gagtggcttc cggagctaac gcgttaagtc gaccgcctgg
 781 ggagtacggc cgcaaggtta aaactcaaat gaattgacgg gggcccgcac aagcggtgga
 841 gcatgtggtt taattcgatg caacgcgaag aaccttacct actcttgaca tccacggaat
 901 ttggcagaga tgccttagtg ccttcgggaa ccgtgagaca ggtgctgcat ggctgtcgtc
 961 agctcgtgtt gtgaaatgtt gggttaagtc ccgcaacgag cgcaaccctt atcctttgtt
1021 gccagcgatt cggtcgggaa ctcaaaggag actgccggtg ataaaccgga ggaaggtggg
1081 gatgacgtca agtcatcatg gcccttacga gtagggctac acacgtgcta caatggcgca
1141 tacaaagaga agcgacctcg cgagagcaag cggacctcac aaagtgcgtc gtagtccgga
1201 tcggagtctg caactcgact ccgtgaagtc ggaatcgcta gtaatcgtgg atcagaatgc
1261 cacggtgaat acgttcccgg gccttgtaca caccgcccgt cacaccatgg gagtgggttg
1321 caaaagaagt aggtagctta accttcggga gggcgcttac
```

D.2 *Pantoea* sp. C9-1 16S ribosomal RNA gene, partial sequence

GenBank number: FJ611817.1
Downloaded from: http://www.ncbi.nlm.nih.gov/nuccore/FJ611817.1 (27 April 2011)

```
LOCUS       FJ611817     1360 bp    DNA     linear   BCT 24-NOV-2009
DEFINITION  Pantoea sp. C9-1 16S ribosomal RNA gene, partial sequence.
ACCESSION   FJ611817
VERSION     FJ611817.1  GI:223557721
KEYWORDS    .
SOURCE      Pantoea sp. C9-1
ORGANISM  Pantoea sp. C9-1 Bacteria; Proteobacteria; ammaproteobacteria;
Enterobacteriales; Enterobacteriaceae; Pantoea.
REFERENCE   1  (bases 1 to 1360)
AUTHORS   Rezzonico,F., Smits,T.H., Montesinos,E., Frey,J.E. and Duffy,B.
TITLE     Genotypic comparison of Pantoea agglomerans plant and clinical
strains
JOURNAL   BMC Microbiol. 9, 204 (2009)
PUBMED   19772624
REMARK    Publication Status: Online-Only
REFERENCE   2  (bases 1 to 1360)
AUTHORS   Rezzonico,F., Smits,T.H.M., Montesinos,E., Frey,J.E. and Duffy,B.
TITLE     Direct Submission
JOURNAL   Submitted (08-JAN-2009) Plant Protection Division,
AgroscopeChangins-Waedenswil ACW, Schloss, Postfach 185, Waedenswil
8820,Switzerland
FEATURES             Location/Qualifiers
     source          1..1360
                     /organism="Pantoea sp. C9-1"
                     /mol_type="genomic DNA"
                     /strain="C9-1"
                     /db_xref="taxon:593074"
     rRNA            <1..>1360
                     /product="16S ribosomal RNA"
```

```
   1 cggacgggtg agtaatgtct ggggatctgc ccgatagagg gggataacca ctggaaacgg
  61 tggctaatac cgcataacgt cgcaagacca aagaggggga ccttcgggcc tctcactatc
 121 ggatgaaccc agatgggatt agctagtagg cggggtaatg gcccacctag gcgacgatcc
 181 ctagctggtc tgagaggatg accagccaca ctggaactga gacacggtcc agactcytac
 241 gggaggcagc agtggggaat attgcacaat gggcgcaagc ctgatgcagc catgccgcgt
 301 gtatgaagaa grccttcggg ttgtaaagta ctttcagcgg ggaggaaggc grkgcggtta
 361 ataaccgcgt cgattgacgt tacccgcaga agaagcaccg gctaactccg tgccagcasm
 421 cgcggtaata cggagggtgc aagcgttaat cggaattact gggcgtaaag cgcacgcagg
 481 cggtctgtta agtcagatgt gaaatccccg ggcttaacct gggaactgca tttgaaactg
 541 gcaggcttga gtcttgtaga gggggggtaga attccaggtg tagcggtgaa atgcgtagag
 601 atctggagga ataccggtgg cgaaggcggc cccctggaca aagactgacg ctcaggtgcg
 661 aaagcgtggg gagcaaacag gattagatac cctggtagtc cacgccgtaa acgatgtcga
 721 cttggaggtt gttcccttga ggagtggctt ccggagctaa cgcgttaagt cgaccgcctg
 781 gggagtacgg ccgcaaggtt aaaactcaaa tgaattgacg gggcccgca caagcggtgg
 841 agcatgtggt ttaattcgat gcaacgcgaa gaaccttacc tactcttgac atccagagaa
 901 yttrgcagag atgchttrgt gccttcggga actstgagac aggtgctgca tggctgtcgt
 961 cagctcgtgt tgtgaaatgt tgggttaagt cccgcaacga gcgcaaccct tatcctttgt
1021 tgccagcgat tcggtcggga actcaaagga gactgccggt gataaaccgg aggaaggtgg
1081 ggatgacgtc aagtcatcat ggcccttacg agtagggcta cacacgtgct acaatggcgc
1141 atacaaagag aagcgacctc gcgagagcaa gcggacctca caaagtgcgt cgtagtccgg
1201 atcggagtct gcaactcgac tccgtgaagt cggaatcgct agtaatcgtg gatcagaatg
1261 ccacggtgaa tacgttcccg ggccttgtac acaccgcccg tctcaccatg ggagtgggtt
1321 gcaaaagaag taggtagctt aaccttcggg agggcgctac
```

D.3 *Erwinia herbicola* 16S ribosomal RNA gene, partial sequence[7]

GenBank number: U80202.1
Downloaded from : http://www.ncbi.nlm.nih.gov/nuccore/U80202.1 (27 Apr 2011)

```
LOCUS       EHU80202     1454 bp    DNA     linear   BCT 28-OCT-1997
DEFINITION  Erwinia herbicola 16S ribosomal RNA gene, partial sequence.
ACCESSION   U80202
VERSION     U80202.1  GI:2570280
KEYWORDS    .
SOURCE      Pantoea agglomerans
ORGANISM  Pantoea agglomerans
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Pantoea.
REFERENCE   1  (bases 1 to 1454)
AUTHORS   Kwon,S.W., Go,S.J., Kang,H.W., Ryu,J.C. and Jo,J.K.
TITLE     Phylogenetic analysis of Erwinia species based on 16S rRNA gene
sequences
JOURNAL   Int. J. Syst. Bacteriol. 47 (4), 1061-1067 (1997)
PUBMED   9336906
REFERENCE   2  (bases 1 to 1454)
AUTHORS   Kwon,S.W. and Go,S.J.
TITLE     Direct Submission
JOURNAL   Submitted (29-NOV-1996) Molecular Genetics Division, National
Institute of Agricultural Science and Technology, 249 Seodun-dong, Suwon,
Suwon 441-707, Korea
FEATURES             Location/Qualifiers
     source          1..1454
                     /organism="Pantoea agglomerans"
                     /mol_type="genomic DNA"
                     /strain="ATCC 33243"
                     /db_xref="taxon:549"
     rRNA            <1..>1454
                     /product="16S ribosomal RNA"

   1 aacacatgca agtcggacgg tagcacagag agcttgctct cgggtgacga gtggcggacg
  61 ggtgagtaat gtctggggat ctgcccgata gaggggggata accactggaa acggtggcta
```

[7] *E. herbicola* strain 33243.

```
 121 ataccgcata acgtcgcaag accaaagagg gggaccttcg ggcctctcac tatcggatga
 181 acccagatgg gattagctag taggcggggt aatggcccac ctaggcgacg atccctagct
 241 ggtctgagag gatgaccagc cacactggaa ctgagacacg gtccagactc ctacgggagg
 301 cagcagtggg gaatattgca caatgggcgc aagcctgatg cagccatgcc gcgtgtatga
 361 agaaggcctt cgggttgtaa agtactttca gcggggagga aggcgacggg ttaataaccc
 421 tgtcgattga cgttacccgc agaagaagca ccggctaact ccgtgccagc agccgcggta
 481 atacggaggg tgcaagcgtt aatcggaatt actgggcgta aagcgcacgc aggcggtctg
 541 ttaagtcaga tgtgaaatcc ccgggcttaa cctgggaact gcatttgaaa ctggcaggct
 601 tgagtcttgt agaggggggt agaattccag gtgtagcggt gaaatgcgta gagatctgga
 661 ggaataccgg tggcgaaggc ggcccctgg acaaagactg acgctcaggt gcgaaagcgt
 721 ggggagcaaa caggattaga taccctggta gtccacgccg tagacgatgt cgacttggag
 781 gttgttccct tgaggagtgg cttccggagc taacgcgtta agtcgaccgc ctggggagta
 841 cggccgcaag gttaaaactc aaatgaattg acgggggccc gcacaagcgg tggagcatgt
 901 ggtttaattc gatgcaacgc gaagaacctt acctactctt gacatccagc gaacttagca
 961 gagatgcttt ggtgccttcg ggaacgctga gacaggtgct gcatggctgt cgtcagctcg
1021 tgttgtgaaa tgttgggtta agtcccgcaa cgagcgcaac ccttatcctt tgttgccagc
1081 gattcggtcg ggaactcaaa ggagactgcc ggtgataaac cggaggaagg tggggatgac
1141 gtcaagtcat catggccctt acgagtaggg ctacacacgt gctacaatgg cgcatacaaa
1201 gagaagcgac ctcgcgagag caagcggacc tcacaaagtg cgtcgtagtc cggatcggag
1261 tctgcaactc gactccgtga agtcggaatc gctagtaatc gtggatcaga atgccacggt
1321 gaatacgttc ccgggccttg tacacaccgc ccgtcacacc atgggagtgg gttgcaaaag
1381 aagtaggtag cttaaccttc gggagggcgc ttaccacttt gtgattcatg actggggtga
1441 agtcgtaaca aggt
```

List of symbols/abbreviations/acronyms/initialisms

ATCC	American Type Culture Collection
Avg	Average
BG	*Bacillus globigii* (bacillus subtilis var niger)
BSC	Biosafety Cabinet
Ct	Crossing threshold
DNA	Deoxyribonucleic Acid
dNTPs	Deoxyribonucleotide Triphosphate
DPG	Dugway Proving Grounds
DRDC	Defence R&D Canada
EPG	Experimental Proving Ground
gDNA	Genomic DNA
LB	Luria-Bertani
MS2	Male Specific Coliphage 2
NA	Not Applicable
NCBI	National Center for Biotechnology Information
NFW	Nuclease-Free Water
NSN	New Substances Notification
NTC	No Template Control
PBS	Phosphate Buffered Saline
PC	Positive Control
PCR	Polymerase Chain Reaction
PFU	Plaque Forming Unit(S)
RNA	Ribonucleic Acid
RT	Room Temperature
TBE	Tris borate EDTA buffer
TTCP	The Technical Co-Operation Program
WGS	Whole Genome Sequencing

DOCUMENT CONTROL DATA

*Security markings for the title, authors, abstract and keywords must be entered when the document is sensitive

1. ORIGINATOR (Name and address of the organization preparing the document. A DRDC Centre sponsoring a contractor's report, or tasking agency, is entered in Section 8.)

DRDC – Suffield Research Centre
Defence Research and Development Canada
P.O. Box 4000, Station Main
Medicine Hat, Alberta T1A 8K6
Canada

2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.)

CAN UNCLASSIFIED

2b. CONTROLLED GOODS

NON-CONTROLLED GOODS
DMC A

3. TITLE (The document title and sub-title as indicated on the title page.)

Genetic confirmation of DRDC – Suffield Research Centre biological field simulants for the New Substances Notification Program

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

Bader, D.; Fisher, G.; Rowsell, S.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total pages, including Annexes, excluding DCD, covering and verso pages.)	6b. NO. OF REFS (Total references cited.)
June 2018	54	28

7. DOCUMENT CATEGORY (e.g., Scientific Report, Contract Report, Scientific Letter.)

Scientific Report

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

DRDC – Suffield Research Centre
Defence Research and Development Canada
P.O. Box 4000, Station Main
Medicine Hat, Alberta T1A 8K6
Canada

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. DRDC PUBLICATION 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-RDDC-2018-R144	

11a. FUTURE DISTRIBUTION WITHIN CANADA (Approval for further dissemination of the document. Security classification must also be considered.)

Public release

11b. FUTURE DISTRIBUTION OUTSIDE CANADA (Approval for further dissemination of the document. Security classification must also be considered.)

12. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Use semi-colon as a delimiter.)

biological field simulant; PCR; in silico analysis; New Substances Notification Program; Environment Canada; Pantoea agglomerans; Bacillus atrophaeus; MS2 virus; E. coli K-12

13. ABSTRACT (When available in the document, the French version of the abstract must be included here.)

In order to continue to use non-pathogenic organisms as biological warfare agent simulants for open air release, DRDC – Suffield Research Centre was required by Environment Canada to confirm the identities of the simulants employed and submit detailed information packages on each of the simulants in order to comply with the *New Substances Notification Regulations*. The simulant organisms were assessed by Polymerase Chain Reaction (PCR) and in silico analysis. The simulant organisms included *Pantoea agglomerans* (a Gram-negative bacterial simulant formerly named *Erwinia herbicola*), *Bacillus atrophaeus* (a Gram-positive bacterial spore simulant formerly named *Bacillus globigii*) and Male Specific Coliphage 2 (MS2, a viral simulant) including its *Escherichia coli* host. Published polymerase chain reaction (PCR) assays for each of the organisms were reviewed, selected, modified as required, and put into practice. *In silico* analysis was conducted to address Environment Canada's concern regarding the specificity of the *aro*Q assay to detect the *P. agglomerans* simulant strain and not related *Pantoea* strains, including plant pathogens. To address residual prevalence of the simulants prior to new releases, attempts were made to isolate them from DRDC – Suffield Research Centre field study sites where they were previously used. The PCR identification method for *B. atrophaeus* was used to confirm the identity of spore isolates found in environmental samples. The PCR identification methods for the other simulants were not employed as they were not found in these samples. The data and information presented in this report led to the approval of outdoor aerosol release of MS2/*E.coli* and *P. agglomerans* at DRDC – Suffield Research Centre by Environment Canada. *B. atrophaeus* (BG) was also approved, but was not required to go through the NSN submission process, as it was instead placed on the NSN Program Domestic Substances List due to its historical use by DRDC – Suffield Research Centre.

Environnement Canada a demandé au Centre de recherche de Suffield de RDDC de confirmer l'identité des organismes non pathogènes simulant des agents de guerre biologique (« simulants ») qui sont dispersés dans l'atmosphère et de lui envoyer un ensemble de renseignements détaillés sur chacun d'eux afin que le Centre se conforme au *Règlement sur les renseignements concernant les substances nouvelles* et puisse continuer cette pratique. Les simulants ont été évalués par réaction en chaîne de la polymérase (PCR) ou analyse *in silico*. Les organismes simulant trouvés étaient notamment *Pantoea agglomerans* (une bactérie à Gram négatif, auparavant appelée *Erwinia herbicola*), *Bacillus atrophaeus* (une spore bactérienne à Gram positif, auparavant appelée *Bacillus globigii*) et le coliphage mâle spécifique 2 (MS2, un simulant de virus) incluant son hôte *Escherichia coli*. Les essais de réaction en chaîne de la polymérase (PCR) de chacun des organismes ont été étudiés, choisis, modifiés au besoin et mis en pratique. L'analyse *in silico* a été réalisée pour répondre aux préoccupations d'Environnement Canada relatives à la spécificité du test aroQ pour la détection de la souche du simulant *P. agglomerans* plutôt que les souches de *Pantoea* non parentes, y compris les pathogènes végétaux. Pour étudier la question de la prévalence résiduelle des simulants avant une nouvelle dispersion, nous avons tenté de les isoler dans les sites d'étude sur le terrain du centre de recherche de Suffield de RDDC où ils avaient été dispersés antérieurement. La méthode d'identification de *B. atrophaeus* a été utilisée pour confirmer l'identité des isolats de spores trouvés dans les échantillons environnementaux. Nous n'avons pas employé les méthodes d'identification par PCR pour les autres simulants, puisque ceux-ci n'ont pas été retrouvés dans ces échantillons. Les données et les informations présentées dans ce rapport ont mené à l'approbation par Environnement Canada de la dispersion à l'extérieur d'aérosol de MS2/*E. coli* et *P. agglomerans* au Centre de recherche de Suffield de RDDC. Le rejet *B. atrophaeus* (BG) a également été approuvé, mais le processus de divulgation de renseignements concernant les substances nouvelles n'a pas été nécessaire, car cet organisme a plutôt été ajouté à la *Liste intérieure des substances* du Programme des substances nouvelles, à cause de son utilisation de longue date par le Centre de recherche de Suffield de RDDC.