Detection and tracking of a novel genetically-tagged biological simulant in the environment

Peter A. Emanuel1, Patricia E. Buckley1, Tiffany A. Sutton1, Jason M. Edmonds1, Andrew M. Bailey1, Bryan A. Rivers1,5, Michael H. Kim1, William J. Ginley1, Christopher C. Keiser1, Robert W. Doherty1, Frank J. Kragl1, Fiona E. Narayanan1, Sarah E. Katoski1,5, Sari Paikoff2, Samuel P. Leppert3, John B. Strawbridge3, Daniel R. VanReenen3, Sally S. Biberos3, Douglas Moore3, Douglas W. Phillips3, Lisa R. Mingioni3, Ogba Melles3, Daniel G. Ondercin3, Beth Hirsh1, Kendall M Bieschke1, Crystal L Harris1, Kristin M. Omberg4, Vipin K. Rastogi1, Sheila Van Cuyk4, and Henry S. Gibbons1

1Edgewood Chemical Biological Center, 5183 Blackhawk Road APG, MD 21010-5183
2Defense Threat Reduction Agency 8725 John J Kingman Road
Fort Belvoir, VA 22060
3Joint Program Executive Office for Chemical and Biological Defense 5183 Blackhawk Road APG, MD 21010-5183
4Los Alamos National Laboratory PO Box 1663 MS C933 Los Alamos, NM 87545

Running Title: Outdoor tracking of a barcoded anthrax simulant

* Corresponding author

Keywords: Bacillus thuringiensis, Bacillus anthracis, aerosol, re-aerosolization, detection, real-time PCR, air sampling, simulant, genetic barcode, endospore
ABSTRACT

A variant of Bacillus thuringiensis var. kurstaki (Btk) containing a single, stable copy of a uniquely amplifiable DNA oligomer integrated into the genome for tracking the fate of biological agents into the environment was developed. Use of genetically-tagged spores overcomes the ambiguity of discerning the test material from pre-existing environmental flora or from previously released background material. In this study we demonstrate the utility of the genetically “barcoded” simulant in a controlled indoor setting and in an outdoor release. In an ambient breeze tunnel test, spores deposited on tiles were re-aerosolized and detected by real-time PCR at distances of 30 meters from the point of deposition. Real-time PCR signals were inversely correlated with distance from the seeded tiles. An outdoor release of powdered spore simulant at Aberdeen Proving Grounds, Edgewood, Maryland was monitored from a distance by a Light Detection and Ranging (LIDAR) laser. Over a two week period, an array of air sampling units collected samples were analyzed for the presence of viable spores and using barcode-specific real-time PCR assays. Barcoded Btk spores were unambiguously identified on the day of the release, and viable material was recovered in a pattern consistent with the cloud track predicted by prevailing winds and by data tracks provided by the LIDAR system. Finally, the real-time PCR assays successfully differentiated barcoded Btk spores from wild-type spores under field conditions.

INTRODUCTION

The development of sensitive and unequivocal approaches for detecting and tracking highly pathogenic bacteria has traditionally relied upon the use of non-pathogenic spore producing Bacillus species as model organisms or simulants, whose physical and biochemical properties mimic those of the threat agent. Bacillus anthracis (Ba) is a proven bio-threat agent (1, 6, 15-16, 21) due its high virulence and the ability to form hardy and persistent spores, which can persist for decades in certain environments (22). Historically, non-pathogenic spore-forming bacteria such as Bacillus atrophaeus var. globigii (BG) have been used as surrogate organisms to simulate Ba (11, 13). The
physical properties and close genetic relatedness of *Bacillus thuringiensis* var. *kurstaki* (Btk) to *Ba*, most notably the presence of an exosporium which is absent from BG, have led to recent preference for the use of Btk over BG (12). However, the use of BG and Btk in test sites is complicated by the fact that both organisms occur naturally in the environment (see reference (19); see also the excellent review of environmental Btk prevalence by Van Cuyk *et al.* (26)). Btk has a long history of use as a biopesticide, dating back to 1929 studies in the northeastern United States that showed Bt to be effective for insect control against the gypsy moth (18). Today the global use of Btk as a pesticide is widespread with over 13,000 tons produced annually, and the USDA estimates the worldwide market for Btk-based products for forestry and agriculture is greater than $80 million per year (2).

For the purposes of open-air release and tracking of bio-warfare (BW) agent surrogate organisms, the natural occurrence and widespread industrial use of BG and Btk make specific identification of newly disseminated isolates difficult (17, 26). Furthermore, once disseminated, spores can persist in the environment for years to decades (23, 25), complicating analysis of subsequent field release tests. Repeated use of BG spores at military proving grounds has saturated such lands with the accumulation of 75 years of biological test materials (K. Kester, personal communication). Other studies have attempted to develop better BW simulants to replace BG spores. One approach modified polystyrene beads to attach moieties that help track these synthetic products. The ability to detect and track the fate of the material is a valuable attribute, but by building a simulant from a synthetic particle, these materials introduce questions about how well they mimic aerosolization and surface adhesion properties of a real BW release event (9). Another approach (4-5) created plasmid vectors that housed multiple gene based targets, but this simulant strategy was not developed for aerosolization testing purposes.

Recently, a unique genetic “tag” was integrated into the bacterial genome for the purposes of developing a traceable simulant (*Buckley et al., manuscript submitted to AEM*). Using genetic exchange technology for *Bacillus* group species and information gleaned from the newly sequenced genome of ATCC 33679, a well characterized HD-1 biopesticide strain of Btk with an exemplary safety record, (3, 8, 14, 24), we designed...
and introduced a series of 46 base-pair (bp) sequence into specific loci of the genome. These insertions lie within intergenic regions chosen to minimize the potential of disrupting protein coding genes or regulatory sequences which could affect the phenotype of the organism. The 46-bp sequence constitutes a “barcode” which creates a lineage-specific locus that is paired with a chromosomal locus. The design allows for the use of a tailor-made real-time polymerase chain reaction (RT-PCR) assay for specific spore detection and differentiation.

The purpose of creating traceable BW simulant was to provide a better way to validate urban threat models, perform tests in which one can sample and detect multiple strains in parallel from a single dissemination event, and reduce test costs by allowing data to be collected from multiple release tests in parallel. Two clear objectives of this study were: a) to demonstrate the suitability of barcoded Btk (BtkB) spores in tracking spore release in breeze tunnel and in open-air field release; b) to demonstrate re-aerosolization of released spores in controlled environment using tailored PCR assays. The results from the first indoor and outdoor field-tests of barcoded strains correlated with the standard microbiological and PCR detection methods, and these are summarized in this paper.

Materials and Methods

*Large-scale Production of Btk strains for dissemination – Wild-type Btk (BtkW)* (ATCC 33679) and BtkB spores derived from ATCC 33679 (Buckley *et al*, Submitted) were grown in NZ Amine A growth medium consisting of the following in grams per liter (g/L): 10.0 Glucose, 5.0 Casein peptone–type S, 1.0 Yeast extract, 4.0 K2HPO4, 3.0 KH2PO4, 0.134 CaCl2·2H2O, 0.02 FeSO4·7H2O, 0.05 MgSO4·7H2O, 0.023 MnSO4·H2O, 0.02 ZnSO4·7H2O and 0.05 milliliter (mL) Anti-Foam 204 (Sigma-Aldrich) per liter (L) culture volume. All medium ingredients were autoclaved together with the exception of glucose and metals. Glucose filtered through 0.8/0.2 micron (µm) filters was aseptically added to the sterile media. The two metal solutions (CaCl2·2H2O FeSO4·H2O, MgSO4·H2O, MnSO4·H2O, ZnSO4·7H2O) were separately prepared, filtered-sterilized,
and added to prevent precipitation of CaSO₄. One liter seed cultures in a 4 L flask were prepared by inoculating with two vials of frozen stock and incubating (Innova 4300, New Brunswick scientific) at 30°C and 200 revolutions per minute (rpm) until the optical density at 600 nanometers (nm) (Genesys 20, ThermoSpectronic) reached approximately 0.4 optical density units. The 1 L seed cultures were aseptically transferred by a pre-sterilized 2 L transfer bottle into the 100 L fermentor (IF 150, New Brunswick Scientific) containing the pre-sterilized NZ Amine A medium. The operating conditions for the 100 L fermentor were controlled at 240 rpm, air flow of 1 vvm (i.e., air volume per liquid volume per minute), 30°C and pH 7.0 (with 3 M H₃PO₄ and 3 M NaOH). When the percentage of spores (as estimated by phase contrast microscopy (BX51, Olympus)), exceeded 95%, the spore suspension was heat-shocked for one hour at 70°C. After cooling down to 30°C, the spore suspension was diafiltrated (0.2 μm Pellicon TFF, Millipore) to approximately 20 L retentate volume, washed five times with de-ionized water, and concentrated to a final volume of approximately 10 L. The washed and concentrated 10 L spore suspension was spray dried (Niro) and subsequently milled (Sturtevant) with Aerosol 202 (Evonik). The spray dried spore preparation was characterized in terms of their particle size distribution (using a TSI, Inc. Aerodynamic Particle Sizer (APS 3321)), moisture content (Karl Fischer C20 Moisture Analyzer, Mettler-Toledo), and viability (serial dilution and plate counting). The final dried and milled products contained 1.1x10¹¹ (BtkW) and 2.6x10¹¹ (BtkB) CFU/g. Analysis of a representative preparation using a Petroff-Hauser chamber revealed that approximately 90% of total spores were viable prior to dissemination.

**Indoor re-aerosolization** – In order to test the capability of BtkB spores to be re-aerosolized, a test was conducted in a 200 m long Ambient Breeze Tunnel. This tunnel provided a controlled indoor environment that simulated outdoor wind conditions. The experimental setup is illustrated in Figure 1A. Vinyl tiles were cleaned with 70% ethanol/water (v/v) and placed down the length of the Ambient Breeze Tunnel at 3 m intervals to a distance of 30 m feet away from the contaminated tiles. At the front of the Ambient Breeze Tunnel, a row of tiles enclosed in boxes were carefully seeded with 100 mg of the aerosolized BtkB spores. Five of the seeded tiles were set aside and not
subjected to a breeze to serve as the positive control tiles. Midway down the tunnel, a single Dry Filter Unit air sampler (DFU; see below) was also situated to collect any spores that would drift down the tunnel. Prior to initiating the test, five unseeded tiles were sampled as negative controls to ensure that the seeding process did not itself cause spores to drift in the sampling area. To begin the test, fans directed at the contaminated tiles were turned on, creating a breeze that was allowed to blow air across the seeded tiles for a period of ten minutes. The velocity of the breeze was measured to be 3.8 meters per second (m/sec) by anemometers positioned just above the seeded tiles. The fans were then turned off and re-aerosolized spores were allowed to settle overnight. In the morning, each tile was swabbed with a single sterilized 2 inch square non-woven cotton fabric wipe (Dukal corporation). The sponges were wetted with 5 mL of PBS from a 50 mL conical tube and then swabbed across the tile covering the entire tile. The wipe was then folded inside out and used to swab the entire tile again in the other direction. The wipes were placed back into their respective 50 mL tubes and processed for analysis.

Outdoor dissemination and detection of Btk strains

The outdoor dissemination took place at the Edgewood Chemical Biological Center’s M field test range, Aberdeen Proving Ground, MD. The M Field test range is an open field on the Edgewood Peninsula on which a large concrete pad is located to facilitate positioning of test equipment (Figure 2). A test grid consisting of multiple Dry Filter Unit (DFU) collectors and a single Joint Biological Point Detection System (see below) was arranged across the field downwind of the release point. The test plan was designed to release 100 grams (g) (~1.1 x 10^{13} spores) of dried BtkB spores on the first day and continue to sample the air for an additional two days. The layout of the test grid and timeline of dissemination and sampling events are shown in Figure 2. On the third day after the initial release event, a team of two people used commercial blowers for a 15 min period to simulate air turbulence around a 30 square meter area near to the site of the release. The air samples were monitored for an additional day to detect barcoded spores that might have been re-aerosolized by this event. On the eighth day, a 100 g batch of wild-type (BtkW; lacking the bar-code) spores (approximately 2.6 x 10^{13} spores) was
released in the identical manner and the air was monitored. The goal of this second release was to demonstrate the selectivity of the detection methodology.

**Dissemination of Btk powder** – The dissemination platform for dispersing the spore preparations consisted of a Metronics Model 10 Skilblower mounted on the bed of a High Mobility Multipurpose Wheeled Vehicle (HMMWV; AM General LLC, South Bend, IN), better known as the Humvee. The spores were released at a distance of approximately 50 m west of the test grid, primarily targeting the northwest corner of the pad (Figure 2). The Skilblower is a high volume dissemination system specifically for dry powders. It utilizes a high power motor blower to disseminate large volumes of powders to great distances, generates high shear forces for maximum breakup of particles, and incorporates a variable feed rate system to provide controlled amounts of powder from 2 - 40 grams per minute (g/min). The Skilblower was mounted at a height of approximately 1.5 m above the bed of the Humvee at an angle of 35 degrees from horizontal. The Humvee was driven at ~3 km per hour in the North/South line until the dissemination was complete (i.e. until all of the powder was disseminated). The powder was manually fed into the Skilblower to better control the timing of the release. The feeding of the biological material was done at a steady rate by holding the container of powder to the intake of the Skilblower at a 90º angle. The first release occurred on November 30, 2011 at 6:24PM over a 5 min period and consisted of 100 g of Btk spores which was calculated to consist of $3.3 \times 10^{11}$ colonies per gram and had a mass mean diameter of 3.39 µm, and a moisture content of 8.85%. The second release occurred 8 days later at 7:24PM over a 6.5 minute period and consisted of 87 g of Btk spores which was calculated to consist of $1.1 \times 10^{11}$ colonies per gram and had a mass mean diameter of 3.72 µm, and a moisture content of 9.09%.

**Standoff detection of Btk clouds** – To track the emerging Btk cloud during the dissemination, a Joint Biological Standoff Detection System (JBSDS) was fielded during the release. The JBSDS is a vehicle-mounted Light Detection and Ranging (LIDAR) system that uses infrared and ultraviolet lasers to detect aerosol clouds out to five kilometers and discriminate biological aerosol clouds from non-biological aerosol clouds using the intrinsic fluorescence of biomolecules. The system was located approximately
800 m from the M Field concrete pad, and scanned an azimuth of approximately 30°.
The system display was remotely linked to the command module located off the M Filed
test grid. This information allowed the test team to deduce in real time that the aerosol
cloud would follow the predicted model track and continue traveling northwest across
the test grid passing directly over the concrete test pad and the centrally located
collection and detection systems.

Point detection of Btk clouds – Detecting and identifying a cloud as it passes over
a fixed location, such as the central concrete pad on the test field, is referred to as point
detection, and for this purpose the Department of Defense’s Joint Biological Point
Detection System (JBPDS) was fielded during this test. The JBPDS is a fully automated
system capable of providing detection and presumptive identification of potential
biological warfare aerosol agents. The JBPDS utilizes an ultraviolet fluorescence based
detector which samples and counts particles from the ambient air. When the algorithm
detects a concentration of biological particles above the average measured background,
it triggers the system to initiate an aerosol collection using a wetted wall cyclone
impinger into a liquid buffer solution. This liquid is then used to inoculate a flat cassette
which houses ten immunochromatographic assay test strips specific for biological threat
agents. For the purposes of the November 30 test, one of the antibody-based test strips
was assigned to provide real time presumptive identification of the Btk spores.

Dry Filter Sampling – An array of DFU 2000 dry filter unit samplers was fielded to
collect the Btk spore release. The DFU is a high volume biological sample collection
system used by the military to collect air samples in the field. Twenty DFU 2000’s were
positioned across the test grid. The DFU collects 1000 L of air per min and consists of a
high flow air sampling pump that collects airborne particulates on 47 millimeter (mm)
diameter polyester filters (PEF-1) filters from an air intake mast which extends 3 m in
height. Over the course of the 12-day test, the units were run for a single 8 h period at
indicated time points for 12 days, the collection filters steriley replaced, and the test
filters transported back to a lab for PCR analysis and microbial culture using the spiral
plating method.
Microbiological culture analysis – Culture analysis of the DFU and tile samples was performed using an Autoplate 4000 and QCounter (Spiral Biotech, Inc., Norwood MA). Each DFU sample was spiral plated in triplicate. An additional 3 log dilution was prepared for samples expected to have high concentrations of Btk (i.e. the DFUs on the "pad" during the releases). The third serial dilution was also spiral plated in triplicate for these samples. Plates were incubated at 30 °C for ~16 hours. The QCounter was used to count the colonies on the plate according to manufacturer’s instructions.

Sample resuspension and disruption – Following sampling, the DFU filters were transported to the laboratory in 50 mL conical tubes and manually suspended in 10 mL phosphate buffered saline (PBS) with 0.1% Triton X-100. A 1 mL aliquot was transferred from the 50 mL conical to a 2 mL screw top tube containing 0.29 g (+/- 0.18 g) of 0.1 mm glass beads (No. 11079110; BioSpec Products, Bartlesville, OK). Samples were processed in a Mini-BeadBeater-96 (BioSpec Products) for 15 minutes. Processed samples were removed and beads were allowed to settle for 2 minutes. The supernatant was removed, and 625 µL of the resulting liquid were processed using the DNA extraction method described below.

DNA Extraction – A BioMek® FX Laboratory Automation Workstation was employed for DNA extractions. An aliquot of 625 µL of a sample was added to 625 µL Lysis Buffer (Blood), (MagneSil ONE Fixed Yield Blood Genomic System Cat. MD1392, Promega, Madison, WI) and shaken at 900 rpm for 1 minute, followed by incubation at room temperature for 10 minutes. MagneSil™ Red (60 µL diluted 2:5; Cat. A1641, Promega, Madison, WI) was added and shaken at 900 rpm for 1 minute, followed by the addition of 250 µL isopropanol and mixing by pipetting on the Biomek® for 5 minutes. The MagneSil™ particles were removed from the solution in two half-volume steps by 1 minute incubations using the MagnaBot® (Cat. V8151, Promega, Madison, WI). The MagneSil™ particles were washed with an additional 250 µL of Lysis Buffer (Blood), shaken at 1500 rpm for 1 minute, and recovered with the MagnaBot® for 1 minute. The MagneSil™ particles were washed with 70% ethanol (200 µL 70% ethanol, shaken at 1500 rpm for 1 minute, recovered with MagnaBot® Magnetic
Separation Device for 1 minute) three times. Particles were dried for 10-15 min at room temperature. Purified DNA was eluted into water (150 µL water, incubated at room temperature for 1 minute, shaken at 1500 rpm for 1 minute, heated at 65°C for 3 minutes, shaken at 1500 rpm for 1 minute, and MagnaBot® for 1 minute) and transferred into a shallow 96-well plate.

Real-time PCR – Amplification, data acquisition, and data analysis were carried out on an Applied Biosystems model 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA). Primers specific for the Common Tag of the two-component Barcode module were utilized from Buckley et al. (submitted). PCR reactions were performed in 50 µL volumes in 96-well optical PCR plates (Cat. 4306737; Applied Biosystems, Foster City, CA). Each reaction was set up using SYBR® Green PCR Master Mix (Cat. 4309155; Applied Biosystems, Foster City, CA), 0.25 µM forward and 1 µM reverse primer, Nuclease-free sterile water, and 1 µL extracted DNA product. The thermocycler conditions were as follows: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s, and 60°C for 1 min, followed by a disassociation stage of 95°C for 15 s, 60°C for 15 seconds, and 95°C for 15 s. For detection of all Btk strains, a non-specific TaqMan® fast block RT-PCR assay for Btk was purchased from the Department of Defense Critical Reagents Program (http://www.jpeocbd.osd.mil/packs/Default.aspx?pg=1205). Btk PCR reactions were performed in 20 µL volumes in 384-well optical PCR plates (Cat. 4309849; Applied Biosystems, Foster City, CA). Each reaction was set up using 14.75 µL of master mix, 0.25 µL of Taq polymerase, and 5 µL of extracted DNA product. The thermocycler conditions were as follows: 50°C for 2 min, 95°C for 20 s, 40 cycles of 95°C for 1 s, and 60°C for 20 s. Analysis for all assays was performed using Sequence Detection Software v.2.3.

RESULTS AND DISCUSSION

Demonstration of the utility of the genetically-tagged spores for re-aerosolization in controlled environment was the key objective of this study. Figure 1 summarizes the results of the detection for the re-aerosolized spores in the Ambient Breeze Tunnel.
based on the distance relative to the contaminated tiles. The data show that deposited spores were re-aerosolized and deposited up to 100 feet away following simulation of light wind. The average cycle threshold (CT) values increased with distance from the seeded tiles (Figure 1B) and tracked with the counts of viable spores (Figure 1C). The average CT values increased with increasing distance from the site of seeded tiles, indicating a decrease in concentration of re-deposited spores.

To demonstrate utility of BtKB spores in open-air release studies, an outdoor dissemination of both BtKB and BtKW spores was followed by a suite of detection equipment arrayed over a test grid (Figure 2) to visualize the plume and collect samples downwind of a line-release for each spore type. A JBPDS unit positioned on the northwest corner of the concrete pad at the test field immediately visualized the released spores. Forty seconds after the long distance JPSDS visualized the emerging cloud (Figure 2, inset), the spore cloud passed over the concrete test pad as predicted. The JBPDS was set to register a positive response if biological particulate matter breached a threshold pre-determined for the system. The unit triggered and radioed an alarm to the command center with maximum counts indicative of a dense cloud based on prior experience (data not shown). The immunochromatographic assays in the JBPDS provided a presumptive positive identification that Btk spores were present. The unit then archived a liquid sample which was later recovered and tested to prove that it did contain the uniquely tagged spores (data not shown).

The results for the air sampling immediately following each of the release events are summarized in Figure 3 (as analyzed using DFU filter units). The CT values and CFU data are available as Table S1 in the Supplementary Material. The difference between the average CT values obtained for the barcode PCR assay and the Btk assay are probably due to the location of the PCR target on either chromosome (Barcode) or a plasmid that is present in higher relative copy numbers (Btk). These results show unambiguous detection of BtkW and BtkB genome in both challenges by RT-PCR. The signals obtained by RT-PCR correlated with the total amount of viable material recovered from the DFUs. While both RT-PCR assays registered positive results in both challenges, the overall signal strength for the BtkB assay was much less than the BtkW assay during the second challenge, with only one of the assays registering a CT
value lower than 37, in contrast to the low CT values observed for the generic Btk PCR assay.

We monitored the ability to recover viable spores and generate RT-PCR signals over time following each release by activating each DFU for an 8-hour period over subsequent days (Figure 4). While the RT-PCR signals rapidly dropped to undetectable levels, viable material was recovered, though at levels near or below the detection limits for each of the assays. The complete CT values and detection data are found in the Supplementary Material. Immediately following the release, large amounts of viable material producing robust RT-PCR signals were recovered from several downwind DFU’s (Figures 3, 4A). The number of viable material dropped rapidly following each release, with an immediate 3-log drop on the day following the release and approximately 1-log decrease per day thereafter. In contrast, very few, if any, viable spores were recovered from DFUs not directly downwind of the release point (Figure 3, 4B). DFU-15 registered a relatively large increase in culturable material on 5-December following an attempt to re-aerosolize deposited BtkB from the field (see below). This material exhibited colony morphologies that were easily distinguishable from Btk (data not shown) and therefore probably represented spores/cells of the environmental microbiome.

The use of the leaf blowers to disturb settled spores on the high grass of the test field on day 3 following the release did not register any additional PCR detections. The failure of the leaf blowers to re-aerosolize the spores is not surprising given that the weather during the period of this test was not conducive to measurement of re-aerosolization. The high humidity and noticeable amounts of early morning dew resulted in significant wetting of the grassy areas surrounding the test field each morning. High humidity and surface moisture are known to reduce re-aerosolization; both serve to increase adhesion between particles and the surface (7, 10). Furthermore, wind speeds during the initial spore release were higher than expected; averaging 17 kph with occasional gusts of wind up to 32 kph. Given the milled dry powder released in this test had a mass mean diameter of 3.39 µm, the high wind speeds on the evening of the release likely resulted in the production of a long, sparse
initial deposition area. While the amount of material re-aerosolized does not scale
directly with surface concentration, the concentration of re-aerosolized material at any
given location is a function of the surrounding surface concentration (20). If spores
were re-aerosolized in this experiment, the resultant concentrations were below the
detection limits of the sampling techniques.

Surprisingly, on the eighth day of the outdoor field test, several collectors yielded
weak positive signals for the barcode assay following the release of *Btk*-*W* spores, which
had been meant as negative control release (Figures 3, 4A). We examined the
possibility that the low-level *Btk* signal observed in the *Btk*-*W* release may have
resulted from cross-reactivity of the PCR assay itself. Our laboratory results (Buckley et
al., submitted), showed no cross-reactivity between wild-type and barcoded strains
under the assay conditions utilized in this study, which used 1 ng of purified genomic
dNA and yielded CT values comparable to those obtained from samples derived from
the *Btk* release. Furthermore, genomic DNA extracted from the *Btk*-*W* preparation
utilized in this field test did not yield the *Btk* signal in real-time PCR assays (data not
shown). We therefore believe it to be unlikely that the *Btk* signal observed during the
*Btk*-*W* release was due to cross-reaction of the PCR assay with wild-type DNA.

As no PCR signal was observed for the *Btk* assay in the period between day 3
and 7, and low level PCR activity began immediately following the wild-type release, the
most likely possibility is that barcoded spores may have adsorbed to the intake ports of
the DFU’s during the initial *Btk* release, which then may have been dislodged by the
influx of *Btk*-*W* spores during the second challenge. Because neither the colony counts
nor the PCR detection assay for *Btk*-*W* dropped to completely undetectable levels prior
to the dissemination of *Btk*-*W*, we cannot definitively rule out the residual presence of
*Btk*-*W* spores on surfaces of the detection equipment. Another possibility is that the
surface of the Humvee may have been residually contaminated with *Btk* from the first
test, and that vibration during the test re-aerosolized material adsorbed to the surface of
the Humvee.

Our laboratory results show that the assays should be specific within the range of
CT values observed in this study. However, because the construction of the first
generation of barcoded spore utilized a small section of host genome sequence as a
binding site for the reverse PCR primer, we cannot definitively rule out asymmetric amplification from wild-type DNA leading to a spurious detection result. Therefore, future modifications of the barcoded spore development strategy are underway to reduce the complexity of producing additional variants and to convert the embedded PCR assay from a SYBR® green detection format to a TaqMan® FRET based assay.

In conclusion, the present study and the data summarized here demonstrated that re-aerosolization did occur in the Ambient Breeze Tunnel. The outdoor test also demonstrated the need to thoroughly decontaminate all equipment used in the outdoor field tests and that the current variant of the barcoded spore required changes in its construction. The insertion of genetic barcodes into biological simulants has great advantages for the test and evaluation methodology, and traceable unique biological materials can provide a better way to validate urban threat release models. In addition, the inclusion of test-specific barcodes introduces the possibility of the simultaneous sampling and detection of multiple strains in a single sample collection point. Use of genetically-tagged BW simulants is predicted to simplify tracking and specific detection, critically needed for generating data to support diffusion models and reduce the test costs by allowing data to be collected from multiple tests with minimal interference from pre-existing spore populations.

Acknowledgements – This study was made possible by ECBC internal funding. Creation of the barcoded strains was supported by the Defense Threat Reduction Agency project number CB3654 to H.S.G. and P.E. Opinions presented in this work are those of the authors and do not represent the official policy of the Army, Department of Defense, or the U.S. Government. All information in this report has been cleared for public release.

Figure Legends

Figure 1 – Re-aerosolization in the Ambient Breeze Tunnel Tested by PCR – A) Experimental design: Twenty tiles (one square foot each) were seeded with powdered
barcoded spores (seed and positive controls). Seed tiles were subjected to 10 minutes of 3.8 meters per second wind disturbance for re-aerosolization. The re-aerosolized spores were allowed to settle on clean tiles arrayed at intervals down a 30 meter-long tunnel followed by analysis by culture and PCR. The negative control data represents a composite of clean tiles prior to re-aerosolization and the samples indicated by their distance from the seeded tiles are the average of 5 replicate tiles, with each PCR assay performed in triplicate. The Dry Filter Unit was placed 15 meters from the seeded tiles and was measured before application of the breeze (T=0) and after (DFU following release). B) Results of re-aerosolization test. Main graph – average of all CT values determined for each distance. Results are presented as the average value ± Standard deviation. Inset – heat map of results from each tile. Each PCR assay was performed in triplicate. Results are presented as the average value for each tile. C) Enumeration of viable colonies from each sample. Results are plotted as the average ± standard deviation for all five tiles in each distance range.

Figure 2 – Experimental design of outdoor dissemination test. The map shows a layout of the test range in detail, with a 100 m x 100 m concrete pad as the center point of the experiment. A line-release was carried out from the position indicated to maximize the capture of agent by an array of Dry Filter Unit 2000 (DFU-2000) air samplers that were deployed based on recommendations by a sensor placement tool from Ensco. The location of the Joint Biological Standoff Detection Systems (JBSDS) LIDAR system is indicated. Inset – LIDAR image of progress of BtkB cloud over the test range. Satellite image was obtained from Google Earth (copyright 2012 DigitalGlobe, GeoEye, US Geological Service, USDA Farm Service Agency).

Figure 3 – Detection of Barcoded strains in the field. Following the release, Dry Filter Units were run for 8 hour periods each day over a 12 day period and the processed samples subjected to PCR tests in triplicate. Resuspended material from the filters was also plated on culture medium to enumerate culturable material. Results are shown for the periods immediately following the release of Barcoded (BtkB) and Wild-type (BtkW) material.
Figure 4 – Detection of Btk strains over time. Culture and RT-PCR results are shown from A) DFU-12 and B) DFU-15 (See Figure 3 for locations of the DFUs on the test grid). a – one of three BtkB assays positive. b - one of three BtkW assays positive. Asterisk indicates that colony morphology was distinct from that of Btk.

Supplementary material:
Table S1: Excel spreadsheet showing real-time PCR and viable material results from the outdoor test.


The development of realistic risk models that predict the dissemination, dispersion and persistence of potential biowarfare agents have utilized nonpathogenic surrogate organisms such as *Bacillus atrophaeus* subsp. *globigii* or commercial products such as *Bacillus thuringiensis* subsp. *kurstaki*. Comparison of results from outdoor tests under different conditions requires the use of genetically identical strains; however, the requirement for isogenic strains limits the ability to compare other desirable properties, such as the behavior in the environment of the same strain prepared using different methods. Finally, current methods do not allow long-term studies of persistence or reaerosolization in test sites where simulants are heavily used or in areas where *B. thuringiensis* subsp. *kurstaki* is applied as a biopesticide. To create a set of genetically heterogeneous yet phenotypically distinguishable strains so that variables intrinsic to simulations (e.g., sample preparation) can be varied and the strains can be tested under otherwise identical conditions, we have developed a strategy of introducing small genetic signatures (“barcodes”) into neutral regions of the genome. The barcodes are stable over 300 generations and do not impact in vitro growth or sporulation. Each barcode contains common and specific tags that allow differentiation of marked strains from wild-type strains and from each other. Each tag is paired with specific real-time PCR assays that facilitate discrimination of barcoded strains from wild-type strains and from each other. These uniquely barcoded strains will be valuable tools for research into the environmental fate of released organisms by providing specific artificial detection signatures.

Spores of *Bacillus anthracis*, the causative agent of anthrax, have been successfully weaponized on large scales in at least two historical offensive biological weapons programs (1, 17, 40, 48). *B. anthracis* spores were disseminated through the mail in the well-documented 2001 anthrax attacks (5, 25–26, 38), and were alleged to have been used as a weapon in the former Rhodesia (29, 32). For this reason, *B. anthracis* remains classified as a category A bioweapon. Their physical hardiness, their resistance to heat and environmental insults, and the relative ease with which spores can be refined, milled, and aerosolized without significant loss of viability make *B. anthracis* a significant concern as a potential weapon. Its historical use as a weapon or bioterrorism agent and the substantial potential economic consequences of anthrax reconfirm the need for robust response regimens that are sufficiently robust while minimizing social and economic disruption.

Despite the clear need to acquire knowledge about *B. anthracis* itself, its virulent nature by multiple routes of infection makes the use of the actual agent (or even attenuated derivatives) in outdoor tests impossible. For this reason, initial efforts to develop non-pathogenic bacterial species as simulants focused on *Bacillus atrophaeus* subsp. *globigii*, a relative of *Bacillus subtilis* (12, 18). *B. atrophaeus* subsp. *globigii* has been used for many years as an outdoor simulant of *B. anthracis* (34). However, subsequent research has shown that, while *B. atrophaeus* subsp. *globigii* does mimic many of the properties of *B. anthracis*, it lacks an exosporium and has different thermal-kill properties (7, 11), which decreases its utility as a simulant for *B. anthracis*. The repertoire of *B. atrophaeus* subsp. *globigii* strains in use is quite small and is restricted to a single lineage with very few available polymorphisms that can discriminate between strains, many of which may affect strain and/or spore phenotypes (12).

The limitations of *B. atrophaeus* subsp. *globigii* as a surrogate for *B. anthracis* have prompted several groups to evaluate *Bacillus thuringiensis* subspecies as potential anthrax surrogates (11, 16). Like *B. atrophaeus* subsp. *globigii*, *B. thuringiensis* strains are not known to cause disease in humans, and many strains are available off-the-shelf as biological pesticides for widespread agricultural use in conventional and organic insect pest control (9). Following widespread outdoor applications in pest control scenarios, *Bacillus thuringiensis* subsp. *kurstaki* strains have been recovered from asymptomatic individuals following widespread aerial spray applications over populated areas (45, 46) without any concurrent epidemiological signs of associated disease (22). While *B. thuringiensis* and its pathogenic phylogenetic neighbors *Bacillus cereus* and *B. anthracis* share a highly conserved core genome, the accessory genome or pan-genome is quite variable (28, 36) and consists mainly of plages and plasmids, which encode most of the strain-specific functions that dictate host tropism (e.g., capsule and toxins). The crystalline toxins expressed by *B. thuringiensis* strains are
specific to insects and are not known to affect mammalian hosts. Thus, *B. thuringiensis* spores share many of the important physical and biochemical characteristics of anthrax spores but do not pose a biological hazard to humans. While the use of *B. thuringiensis* as an anthrax simulant is not a novel idea (United Nations inspectors recovered a toxinless strain from a suspected biowarfare facility in Iraq in the late 1990s [8]), it has not yet been widely adopted. The widespread application of *B. thuringiensis* (particularly *B. thuringiensis* subsp. *kurstaki*) as a biopesticide has recently facilitated experimental studies of the persistence and transport of *B. thuringiensis* in the environment (46, 47). While those studies have provided extremely valuable information about the life cycle of deliberately released *B. thuringiensis* spores, the agricultural application of commercial *B. thuringiensis* preparations may not mimic the anticipated aerosol dissemination of an authentic biowarfare agent, confounding the ability to develop realistic models.

Gathering accurate information about organism behavior in the environment requires a combination of robust and reproducible sampling techniques, rigorous methods, and, optimally, a well-characterized input strain. Until now, only very limited numbers of suitable strains existed, limiting the number of possible studies in any given area or time until the recoverable signature returned to background levels. Particularly with persistent spores and in heavily used areas such as the U.S. Army’s Dugway Proving Ground, low-level positive signals could be either authentic or spurious, potentially resulting from reaerosolization of spores left over from previous tests. In fact, the level of residual *B. atrophaeus* subsp. *globigii* spores in the soils at Dugway Proving Ground is as high as 10^5 spores/g soil (K. Omberg, personal communication). The lack of specific signatures for any given strain has made the differentiation of those events impossible.

As a potential solution to this problem, we describe here a new approach to simulant development whereby a stable genetic tag, or “barcode,” is integrated directly into the chromosome of a *B. thuringiensis* subsp. *kurstaki* strain. Each barcode contains two tag modules, one common to all barcoded strains and one specific for each strain. To facilitate the detection and quantitation of each barcoded strain, tag-specific real-time PCR assays that can distinguish the strains from each other, from wild-type strains, and from a panel of near-neighbors and other potential interfering agents are described. We present data on the stability of the barcode during serial transfer and show that the insertion is neutral for *in vitro* growth kinetics. The development of new, specific strains will have a dramatic impact on the methodology of testing and analysis of environmental releases.

### MATERIALS AND METHODS

#### Strains and plasmids

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Description</th>
<th>Source or reference</th>
</tr>
</thead>
</table>
| **Strains**
| *B. thuringiensis* subsp. *kurstaki*
| ATCC 33679 | HD-1 biopesticide strain | ATCC^a |
| | T1B1 | ATCC 33679 ΔpHD1-XO1; barcoded at target 1 with common tag and specific tag 1 | This work |
| | T1B2 | ATCC 33679 ΔpHD1-XO1; barcoded at target 1 with common tag and specific tag 2 | This work |
| | Foray Commercial HD-1 biopesticide product dispersed in Fairfax County, VA | 45 |
| *E. coli*
| SM10 | *E. coli* donor strain | 24 |
| | SCS110 | pSS4333 donor strain | 24 |
| **Plasmids**
| pRP1028 | Allelic exchange vector, turbo-uff, Spc^c | 24 |
| pS54332 | l-SceI expression vector, gfp, Kan^r | 24 |
| pT1T1B | pRP1028 containing target 1 with common tag and specific tag 1 | DNA2.0 |
| pT1T1B2 | pRP1028 containing target 1 with common tag and specific tag 2 | DNA2.0 |

^a ATCC, American Type Culture Collection.

#### Strains and plasmids

The widespread application of *B. thuringiensis* spores, the agricultural application of commercial *B. thuringiensis* preparations may not mimic the anticipated aerosol dissemination of an authentic biowarfare agent, confounding the ability to develop realistic models. Gathering accurate information about organism behavior in the environment requires a combination of robust and reproducible sampling techniques, rigorous methods, and, optimally, a well-characterized input strain. Until now, only very limited numbers of suitable strains existed, limiting the number of possible studies in any given area or time until the recoverable signature returned to background levels. Particularly with persistent spores and in heavily used areas such as the U.S. Army’s Dugway Proving Ground, low-level positive signals could be either authentic or spurious, potentially resulting from reaerosolization of spores left over from previous tests. In fact, the level of residual *B. atrophaeus* subsp. *globigii* spores in the soils at Dugway Proving Ground is as high as 10^5 spores/g soil (K. Omberg, personal communication). The lack of specific signatures for any given strain has made the differentiation of those events impossible.

As a potential solution to this problem, we describe here a new approach to simulant development whereby a stable genetic tag, or “barcode,” is integrated directly into the chromosome of a *B. thuringiensis* subsp. *kurstaki* strain. Each barcode contains two tag modules, one common to all barcoded strains and one specific for each strain. To facilitate the detection and quantitation of each barcoded strain, tag-specific real-time PCR assays that can distinguish the strains from each other, from wild-type strains, and from a panel of near-neighbors and other potential interfering agents are described. We present data on the stability of the barcode during serial transfer and show that the insertion is neutral for *in vitro* growth kinetics. The development of new, specific strains will have a dramatic impact on the methodology of testing and analysis of environmental releases.

#### MATERIALS AND METHODS

| Strains and plasmids | Strains and plasmids utilized in this study are shown in Table 1. Since it was expected that the tagged spore would be used in a broad range of indoor and outdoor test scenarios, strain ATCC 33679, an HD-1 strain (serotype 3a3b) that is registered with the United States Environmental Protection Agency as an approved biopesticide, was selected as the backbone for the barcoding efforts for its outstanding safety record in widespread gypsy moth control efforts, with annual outdoor applications of ~453 metric tons of *B. thuringiensis* subsp. *kurstaki* spores applied over >138,000 Ha in the United States alone with no significant medical issues recorded (44). The ATCC strain was confirmed to be an HD-1 strain of *B. thuringiensis* by comparison of plasmid profiles to previously published work (41) and by whole-genome sequence analysis with *in silico* multilocus sequence typing (MLST), amplified fragment length polymorphism (AFLP), and *cry* gene typing. Unless otherwise indicated, strains were grown on brain heart infusion agar (BHI) containing polyoxin B (50 μg/ml) or kanamycin (250 μg/ml) and either spectinomycin (50 μg/ml) or kanamycin (20 μg/ml). Unless otherwise noted, strains were incubated at 30°C.

#### Identification of a barcode insertion point

Potential insertion sites for the barcodes were identified based on a set of selection rules elaborated in Table 2. Insertion points were identified in the published genome sequence of *B. thuringiensis* subsp. *kurstaki* strains BMB171 (19) and T03a001 (Refseq accession number BC_CM000751.1).Annotations of the BMB171 genome generated in RAST (4) and PATRIC (15) were compared. We also generated a draft genome sequence of ATCC 33679 (M. Krepps, S. Broomall, P. Roth, C. N. Rosenzweig, and H. S. Gibbons, unpublished data) and verified that the genome structure fulfilled the appropriate criteria. Of 294 intergenic regions >500 bp long (see Table S1 in the supplemental material), three potential target insertion points were identified (Table 3) that fulfilled all of the set criteria.
TABLE 2 Selection rules for barcode insertion points

<table>
<thead>
<tr>
<th>Rule</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target region must be located in the chromosome</td>
<td>Maximize stability by incorporation on major replicon</td>
</tr>
<tr>
<td>Insertion point must lie near the midpoint of an intergenic space larger than 500 bp</td>
<td>Minimize disruption of potential coding sequences or regulatory elements</td>
</tr>
<tr>
<td>No annotated genes or potential ORFs in the intergenic space</td>
<td>Minimize disruption of potential coding sequences or regulatory elements</td>
</tr>
<tr>
<td>Must lie between two convergently transcribed genes</td>
<td>Minimize disruption of potential coding sequences or regulatory elements</td>
</tr>
<tr>
<td>No repetitive structure in intergenic space</td>
<td>Facilitate synthesis and cloning of constructs and minimize potential issues with homologous recombination</td>
</tr>
<tr>
<td>No identical repetitive elements &gt;200 bp in size within 10,000 bp</td>
<td>Minimize potential loss by deletion via homologous recombination between repeat elements (e.g., insertion sequences)</td>
</tr>
<tr>
<td>Target must be present in commercial B. thuringiensis subsp. kurstaki isolate</td>
<td>Maximize likelihood of success in selected target strain</td>
</tr>
<tr>
<td>Target must be intact and consistently annotated in two or more available B. thuringiensis subsp. kurstaki sequences and in ATCC 33679 draft</td>
<td>Maximize versatility and adaptability of barcode targeting vectors to different strains</td>
</tr>
</tbody>
</table>

Barcode module design. We appropriated a set of published 20-bp tags previously used in signature-tagged mutagenesis studies of pooled yeast strains (33). Tags were individually screened against the B. thuringiensis subsp. kurstaki genome sequences to eliminate sequences that had homology to any portion of the B. thuringiensis subsp. kurstaki chromosome. One tag was adopted as a common tag to be shared among multiple strains, while the others were used as strain-specific tags (S1, S2, etc.). The tags were flanked by an EcoRI restriction site to facilitate screening of recombinant strains. Figure 1 shows the general features of a barcode module and the design of associated real-time PCR assays.

Barcode insertion. Barcodes flanked by ~750 bp of chromosomal DNA sequence were generated synthetically (DNA2.0, Menlo Park, CA) and cloned into pRP1028, which was delivered by a protocol adapted from the work of Jances and Stibitz (24). The resulting plasmids were delivered by biparental mating into ATCC 33679. Replication of pRP1028 was suppressed by maintaining strains at 37°C. Strains which had integrated the plasmids by homologous recombination were selected on spectinomycin plates. Fluorescence of intact colonies due to the turbo-rfp on pRP1028 was checked by transillumination. The I-SceI-expressing plasmid pSS4333 was delivered by triparental mating into the integrant strains. Green-fluorescing Spc+ colonies were screened for the presence of the barcode by PCR amplification and EcoRI digestion of the target locus. pSS4333 was cured by serial transfer on solid media in the absence of selection. The curing of the plasmids was verified by checking the strains for the absence of red or green fluorescence and by the lack of PCR amplification of plasmid-borne antibiotic resistance genes spc and kan.

Real-time PCR assays. Primers and concentrations used for the strain construction, verification, and detection of the barcodes are shown in Table S2 in the supplemental material. Barcodes were detected by real-time SYBR green PCR assays in 20-µl volumes in 384-well optical PCR plates. Amplification, data acquisition, and data analysis were carried out on an Applied Biosystems model 7900HT sequence detection system (Applied Biosystems, Foster City, CA). The barcode reactions were set up using SYBR green PCR master mix (catalog no. 4309155; Applied Biosystems, Foster City, CA), forward and reverse primers, nuclease-free sterile water, and 1 µl extracted DNA product. The thermocycler conditions for the common tag and barcode 2 were as follows: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s, and 60°C for 1 min, followed by a disassociation stage of 95°C for 15 s, 60°C for 15 s, and 95°C for 15 s. The barcode 1 thermocycler conditions were set up similarly to the program above, with the following exception: annealing was at 55°C for 15 s (instead of 60°C for 15 s). The linear range for each reaction was determined by developing a standard curve for eight 10-fold serial dilutions of the corresponding genomic DNA. The efficiency of each reaction was calculated from the resulting graphs.

Barcode stability. A single starter culture of each strain was grown in 5 ml of medium. Three independent 50-ml cultures of each B. thuringiensis subsp. kurstaki strain (wild type and two strains containing different specific tags in target 1) were grown in BHI medium in shaking flasks at 30°C. Each day, at approximately the same time, cultures were diluted 1:1,000 in fresh medium. The process was repeated for 5 days, after which the cultures were allowed to incubate at 30°C for 3 days in order to induce sporulation. This cycle was repeated each week for 6 weeks, representing approximately 300 doublings. Where applicable, growth was monitored by optical density at 600 nm (OD600).

Comparative growth of barcoded strains. Barcoded strains were grown either in a 20-liter fermentor or in parallel flask cultures. For parallel flask cultures, samples were withdrawn periodically for determination of the OD600. For growth in the fermentors, wild-type and barcoded strains were grown in 20-liter volumes of NZ-Amine A medium in Micros 30 fermentors (New Brunswick Scientific, Enfield, CT). Starter cultures of 500 ml were grown in 2-liter flasks in a shaking incubator until the OD600 reached ~0.4. Seed cultures were aseptically transferred into the Micros 30 fermentor (with a 20-liter working volume) containing the NZ-Amine A medium. The operating conditions for the Micros 30- liter fermentor were controlled, with an agitation speed of 300 rpm and an airflow of one air volume per liquid volume per minute at 30°C and pH 7.0. Dissolved oxygen (percent saturation) and optical density (600 nm) were monitored using an in-line probe and by periodic sampling, respectively.

TABLE 3 Potential barcode insertion points

<table>
<thead>
<tr>
<th>Target</th>
<th>Intergenic locus (BMB171 coordinates)</th>
<th>Flanking gene and product</th>
<th>5'</th>
<th>3'</th>
<th>Intergenic gap size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>882263–882815</td>
<td>BMB171_C0768 acyl-coenzyme A synthetase</td>
<td>BMB171_C0769, hypothetical protein</td>
<td>552</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1533619–1534178</td>
<td>BMB171_C1412 thiosulfate sulfurtransferase</td>
<td>VBI_Bachui4800_1517,2 hypothetical protein</td>
<td>559</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2786602–2787119</td>
<td>BMB171_2615 hypothetical protein</td>
<td>BMB171_C2616, alkane-sulfonate monoxygenase (ssdD)</td>
<td>517</td>
<td></td>
</tr>
</tbody>
</table>

* PATRIC annotation; no NCBI locus tag called for this gene.
Genomic characterization of the barcoded strain. A 454 shotgun draft sequence of the barcoded strain was generated by standard methods using the 454 Titanium package (Roche/454, Branford, CT). Reads were mapped to the scaffolds from the parent strain that had been generated from melded 454 shotgun and paired-end libraries using Newbler v2.6 (Krepps et al., unpublished). Reads were mapped to the parent strain using the mapping algorithm in Genomics Workbench from CLC Bio (Aarhus, Denmark) using the default parameters. Regions of low or absent sequence coverage were identified, and deletion endpoints, where applicable, were identified by manual inspection of the mapped data.

RESULTS

Integration of barcodes. We selected *B. thuringiensis* subsp. *kurstaki* strain ATCC 33679, a prototypical HD-1 strain of *B. thuringiensis* subsp. *kurstaki*, for our barcoding efforts. Our selection was guided by the long history of the use of HD-1 strains as EPA-approved biopesticides, dating as far back as 1961 (23). *B. thuringiensis* subsp. *kurstaki* HD-1 is the active ingredient in Foray, a commercial *B. thuringiensis* subsp. *kurstaki* product. Target sequences were identified by PCR in both ATCC 33679 and a sample of Foray (Fig. 2A). Barcoded target constructs were synthesized and successfully integrated at two of the three identified loci (targets 1 and 2). PCR amplification of target sequences revealed the presence only of EcoRI-digestible product and not the parent product, indicating successful replacement of the parental allele (Fig. 2B). Similar results were obtained for target 2 (data not shown). Repeated attempts to integrate a barcode into target 3 were unsuccessful for reasons that are not clear at this time.

Real-time PCR detection assay. To allow easy detection of the barcoded strains, real-time SYBR green PCR assays specific to the common and specific tags were developed. The assay directed at the common tag recognized both barcoded strains, whereas the assays directed at the specific tags recognized only their cognate strains. Because one of the primers for each sequence is derived from endogenous genetic material, careful control over primer concentrations and PCR amplification conditions was found to be critical to avoid spurious false-positive signals due to asymmetric amplification. After careful optimization, none of the assays directed at the barcoded strains recognized the wild-type strain.
Figure 3A shows representative real-time PCR assay traces and standard curves (Fig. 3C) for each assay. Table 4 lists the linear range and limit of detection of each assay, along with its calculated efficiency. Based on the 11.2-Mbp estimated genome size obtained from the Newbler de novo assembly, which weights the sequence coverage of each element rather than the total size of the assembly, the detection limit for each assay is approximately 8 to 80 genome copies. The differences in the limit of detection between the assays are most likely attributable to the differences in GC content between the chromosomal primer binding sites.

**TABLE 4 Sensitivity of PCR assays for *B. thuringiensis* subsp. *kurstaki* strains**

<table>
<thead>
<tr>
<th>Assay</th>
<th>Efficiency (%)</th>
<th>Linearity</th>
<th>Estimated LOD&lt;sup&gt;a&lt;/sup&gt; (no. of genome copies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common tag</td>
<td>75</td>
<td>1 pg to 10 ng</td>
<td>83</td>
</tr>
<tr>
<td>Specific tag 1</td>
<td>72</td>
<td>1 pg to 10 ng</td>
<td>83</td>
</tr>
<tr>
<td>Specific tag 2</td>
<td>77</td>
<td>100 fg to 10 ng</td>
<td>8.3</td>
</tr>
</tbody>
</table>

<sup>a</sup> LOD, limit of detection.

**Near-neighbor panel screens: inclusivity and exclusivity.** Using the real-time SYBR green PCR assays discussed above, we tested the barcode PCR assays for specificity against the barcoded strains themselves, their wild-type parent strains, and a selection of related and unrelated bacterial strains (Table 5). The barcode assays were specific for their cognate targets and did not yield amplicons with the unmarked and near-neighbor strains. When nonspecific amplification was observed, these amplicons produced dramatically higher threshold cycle ($C_T$) values and no dis-
TABLE 5 Specificity of real-time PCR assays for B. thuringiensis subsp. kurstaki strains

<table>
<thead>
<tr>
<th>Strain or material tested</th>
<th>Specific tag 1</th>
<th>Specific tag 2</th>
<th>Specific tag 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. thuringiensis subsp. kurstaki</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>B. thuringiensis subsp. kurstaki T1B1</td>
<td>4/4 (22.0)</td>
<td>4/4 (25.2)</td>
<td>0/4</td>
</tr>
<tr>
<td>B. thuringiensis subsp. kurstaki T1B2</td>
<td>4/4 (27.2)</td>
<td>0/4</td>
<td>4/4 (22.0)</td>
</tr>
<tr>
<td>Bacillus anthracis VNR-Δ1</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>B. anthracis Ames</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>B. anthracis NNR-Δ1</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>B. anthracis ΔSterne</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>B. thuringiensis subsp. israelensis ATCC 35646</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>Bacillus cereus HER1414</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>B. subtilis ATCC 27370</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>B. atrophaeus subsp. globigii</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa PAO-1</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>Streptococcus pyogenes ATCC 12384</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>Bordetella pertussis ATCC 9797</td>
<td>0/4</td>
<td>3/4 (39.0)</td>
<td>0/4</td>
</tr>
<tr>
<td>Salmonella enterica serovar Typhimurium ATCC 14028</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>Escherichia coli ATCC 43985</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>Human placental DNA</td>
<td>0/4</td>
<td>1/4 (39.8)</td>
<td>0/4</td>
</tr>
<tr>
<td>Escherichia coli O157:H7</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>Francisella tularensis SHU4</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
<tr>
<td>Yersinia pestis HARBN35</td>
<td>0/4</td>
<td>0/4</td>
<td>0/4</td>
</tr>
</tbody>
</table>

* Number of samples crossing threshold per number of replicates tested. One nanogram of genomic DNA was tested per replicate. Values in parentheses are average Ct values.

** Genome resequencing.** To identify any other potential genetic alterations to the barcoded strains and to verify unambiguously the location and uniqueness of the barcode insertion, we resequenced one of the barcoded strains and mapped the data onto a Newbler assembly of ATCC 33679. The barcode insertion points were evident at the specified locus (Fig. 5); our resequencing data indicated that approximately 350 kb of genetic material had been lost at some point during the strain construction process. Most of the deleted material corresponded to three of the scaffolds regions annotated as plasmids. Bioinformatic analysis of the annotated features generated in RAST and comparison with the deleted genes (see Table S3 in the supplemental material) revealed that most of the deleted material was likely one or more of the many plasmids present in B. thuringiensis subsp. kurstaki and B. cereus strains (3, 21, 35, 37). The genes lost included many homologues of genes on B. anthracis plasmid pXO1 (35). These plasmids, including pXO1 itself from B. anthracis, are readily cured during growth at higher temperatures (2, 39, 43), and given the requirement of prolonged 37°C incubation to suppress plasmid replication during the homologous recombination phase of strain construction, the loss of such material is not surprising.

**DISCUSSION**

We have successfully introduced small genetic barcodes—short, specific identifying signatures—into the genome of B. thuringiensis subsp. kurstaki and coupled the integration of those signatures to specific real-time PCR detection assays directed toward those barcodes. Our work differs from the widely used "signature tags" that uniquely identify transposon insertions and track abundance of individual mutant pools in large populations (20, 33), in that we aim to tag a single locus in multiple isolates with multiple stable chromosomal tags. Our work is similar in intent to the efforts in the synthetic biology community, which added specific “watermarks” to differentiate synthetic genomes from their natural counterparts (13, 14), in that it seeks to incorporate simple, neutral signatures into the chromosome as means of uniquely marking a strain. In fact, our efforts expand upon the idea of watermarking strains by developing specific detection assays based on real-time PCR. Indeed, as described in the accompanying article, these assays allow the detection and differentiation of barcoded B. thuringiensis subsp. kurstaki strains both in the laboratory and in the field (10).

The ability to assign a specific marker to a strain used in field...
FIG 5 Whole-genome sequencing to verify barcode insertion. 454 shotgun sequencing reads were mapped onto the draft genome sequences of the wild-type (top) and in silico-modified barcoded (bottom) genomes. Perfect matches to reference sequence (top row in each panel) are indicated by dark letters; imperfect matches are indicated by lightly shaded letters.
release studies has additional potential advantages, particularly with viable bioinsecticides. Like simulant releases in heavily used proving ground areas, attempts to attribute infections with Bacillus spp. in areas of widespread B. thuringiensis subsp. kurstaki application to the B. thuringiensis subsp. kurstaki serotype actually applied to the treated area have been confounded by ubiquitous environmental B. thuringiensis isolates that cannot always be unambiguously distinguished from the biopesticide strain (reference 45 and references therein). The barcodes endow the new simulant strains with a unique signature that can allow us to definitively exclude simulant B. thuringiensis subsp. kurstaki strains as causative agents of suspected B. thuringiensis infections that might coincide with simulant releases, and if adopted by commercial pesticide manufacturers, these barcodes could serve as exclusionary markers for the biopesticide strains. Furthermore, insertion of the tag into the chromosome minimizes chances of transfer to other strains or species, as the rates of transfer of chromosomal loci between B. thuringiensis strains are quite low (2).

We acknowledge that the process used to insert the barcode may have caused the curing of one or more plasmids and/or the loss of chromosomal material. In previous years, the ability to perform post hoc genomic characterization of mutagenized strains was cost prohibitive. However, modern whole-genome sequencing allows the detection of such events and allows the precise inventory of genomic content of product strains. The loss of genetic material during in vitro culture of Bacillus strains is not surprising—plasmids are often unstable at high temperatures, and the strains carry a complement of genetic material that may be unfavorable during in vitro growth in rich medium. Recombination of T1B2 into the chromosome during the first step of homologous recombination occurred with much lower frequency than that of T1B1 (data not shown); in fact, the only successful integrant obtained was the deletion construct described here. In contrast, multiple successful integrants were obtained for T1B1. Together, these results suggest that one or more elements of T1B2 may be incompatible with a plasmid-encoded functionality, most likely putative restriction endonuclease encoded within the 400 kb of deleted material. While our barcode does not contain any obvious candidates for a restriction endonuclease recognition site, we cannot exclude the possibility that it may be sensitive to an endonuclease activity that is specific to a sequence motif present only in barcode 2; the most likely candidate at this time is the GATC consensus Dam methylation site present in barcode 2 (Fig. 1). While the effect of the loss of ~400 kb of genetic material (see Table S3 in the supplemental material) in our strains is not immediately evident, inferences might be gained from studies of plasmid loss in B. anthracis. In particular, loss of pXO1 from B. anthracis strains is associated with numerous phenotypic changes, including changes to sporulation kinetics, nutritional requirements, and phage sensitivity (42). Our strain also had lost a suite of genes involved in the biosynthesis of zwittermicin, a biologically active compound that, among other activities, potentiates the activity of crystal toxins in insect hosts (6, 27). The loss of this gene cluster containing large polyketide synthase modules is reminiscent of the early loss of surfactin biosynthesis during the domestication of B. subtilis and B. atrophaeus subsp. globigii strains (12, 30, 31). Based on our phenotypic analysis, the effects of the loss on in vitro growth, colony morphology, and sporulation of the barcoded strains appear to have been minimal. We are attempting to recreate the barcoded strains to retain and/or restore as much of the full complement of genes as possible.

The ability to differentiate two tagged strains based on the site-specific integration of specific genetic tags will allow controlled studies in situations where variables previously could not be controlled. For example, it is anticipated that two strains could be prepared or disseminated using different methods, released and collected simultaneously under identical environmental conditions, and then tracked independently in a single set of samples. Our data indicate that simultaneous detection and quantitation may be possible in mixtures containing wild-type and barcoded strains. Alternatively, a test area could be reused quickly without having to wait for the detection signals to return to background levels.

We believe that our barcoding strategy will be generally applicable to genetically tractable microorganisms, although the specific barcode sequences, the genetic tools required to deliver barcodes, and the actual insertion points for the barcodes themselves will differ from organism to organism. These variations will be based on the overall and local genetic structure of the target organism. We are currently automating the bioinformatic identification of barcode insertion points and the design of barcode modules to maximize specificity, sensitivity, and selectivity across a broad range of potential target organisms.

ACKNOWLEDGMENTS

This work was supported by the Defense Threat Reduction Agency (DTRA) project number CB3654 to H.S.G. Sequencing of the B. thuringiensis subsp. kurstaki isolate was supported by DTRA project number CB2847 to H.S.G., C.N.R., and E.W.S.

We thank Jason Edmonds and Vipin Rastogi for critical reading of the manuscript.

The opinions stated in this article are those of the authors and do not represent the official policy of the U.S. Army, Department of Defense, or the Government of the United States. Information in this report is cleared for public release.

REFERENCES


