Strain-Specific Single-Nucleotide Polymorphism Assays for the Bacillus anthracis Ames Strain†

Matthew N. Van Ert,¹ W. Ryan Easterday,¹ Tatum S. Simonson,¹ Jana M. U’Ren,¹ Talima Pearson,¹ Leo J. Kenefic,¹ Joseph D. Busch,¹ Lynn Y. Huynh,¹ Megan Dukerich,¹ Carla B. Trim,¹ Jodi Beaudry,¹ Amy Welty-Bernard,¹ Timothy Read,³ Claire M. Fraser,³ Jacques Ravel,³ and Paul Keim¹,²*

Northern Arizona University, Department of Biological Sciences, Box 5640, Flagstaff, Arizona 86011; Translational Genomics Research Institute, Pathogen Genomics Division, 445 N. Fifth Street, Phoenix, Arizona 85004; and The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20723

Received 15 June 2006/Returned for modification 22 July 2006/Accepted 7 October 2006

Highly precise diagnostics and forensic assays can be developed through a combination of evolutionary analysis and the exhaustive examination of genomic sequences. In Bacillus anthracis, whole-genome sequencing efforts revealed ca. 3,500 single-nucleotide polymorphisms (SNPs) among eight different strains and evolutionary analysis provides the identification of canonical SNPs. We have previously shown that SNPs are highly evolutionarily stable, and the clonal nature of B. anthracis makes them ideal signatures for subtyping this pathogen. Here we identified SNPs that define the lineage of B. anthracis that contains the Ames strain, the strain used in the 2001 bioterrorist attacks in the United States. Sequencing and real-time PCR were used to validate these SNPs across B. anthracis strains, including (i) 88 globally and genetically diverse isolates; (ii) isolates that were shown to be genetic relatives of the Ames strain by multiple-locus variable number tandem repeat analysis (MLVA); and (iii) several different lab stocks of the Ames strain, including a clinical isolate from the 2001 letter attack. Six SNPs were found to be highly specific for the Ames strain; four on the chromosome, one on the pX01 plasmid, and one on the pX02 plasmid. All six SNPs differentiated the B. anthracis Ames strain from the 88 unique B. anthracis strains, while five of the six separated Ames from its close genetic relatives. The use of these SNPs coupled with real-time PCR allows specific and sensitive (<100 fg of template DNA) identification of the Ames strain. This evolutionary and genomics-based approach provides an effective means for the discovery of strain-specific SNPs in B. anthracis.

The 2001 anthrax letter attack illustrated the “real-world” efficacy of Bacillus anthracis as a bioterror agent. Forensic and epidemiological analysis of clinical samples and weaponized spores from the letter attack included the identification of the B. anthracis strain as Ames (3, 8, 17). This was initially accomplished using multiple-locus variable number tandem repeat analysis (MLVA) (3, 6, 8), which was one of the few technical approaches possible for this highly monomorphic pathogen (5). Identifying the strain and establishing its identity in attack locations were important in linking the dispersed anthrax cases and suggesting a possible source for the weaponized material. During the anthrax attack crisis, diagnostic speed, specificity, and sensitivity were often limiting factors, necessitating extraordinary efforts by public health officials and forensic labs (3). Further advancement of molecular diagnostics allows for more-efficient responses in disease outbreaks, whether natural or bioterrorist mediated.

Single-nucleotide polymorphisms (SNPs) are increasingly recognized as important markers for detecting and subtyping bacterial pathogens, including B. anthracis (1, 2, 4, 8, 16, 18, 19). Recent comparative full-genome sequencing allowed the discovery of about 3,500 SNPs among eight strains of B. anthracis (15, 17; J. Ravel, unpublished data) and represents a valuable resource for developing diagnostic signatures for this pathogen. SNPs are particularly attractive markers for subtyping efforts since they are evolutionarily stable (8, 15) and are amenable to high-throughput detection methods, such as real-time PCR, pyrosequencing, and mass spectrometry (1, 2, 4, 10, 20, 22, 23).

As binary markers, SNPs would seem to have limited subtyping power, requiring a large number of SNPs to be interrogated for subtyping purposes. However, researchers have demonstrated that a small number of SNPs can be used to efficiently define genetic groups. For example, Moorhead et al. (11) used a phylogenetic approach to identify a small number of diagnostic SNPs in the sigB gene that partitioned strains of Listeria monocytogenes into three previously described clonal lineages. Keim et al. (8) expanded upon this concept and proposed the idea of a “canonical SNP” (canSNP), an SNP that can be used to efficiently identify a point in the evolutionary history of a species. Mapping SNPs on a genetic population structure allows the identification of canSNPs that can be used diagnostically on a broad scale to define major genetic lineages in a species or, more narrowly, to define specific strains (8). A more recent publication used a similar approach to identify a streamlined set of SNPs that divide methicillin-resistant Staphylococcus aureus isolates into genetic groups that are consistent with its population structure (18).

* Corresponding author. Mailing address: Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ 86011-5640. Phone: (928) 523-1078. Fax: (928) 523-0639. E-mail: Paul.Keim@nau.edu.
† Supplemental material for this article may be found at http://jcm.asm.org.
‡ Published ahead of print on 8 November 2006.
In this study, we used a subset of about 1,000 previously published SNPs (15, 17) to identify canSNPs that effectively define the *B. anthracis* lineage that contains the Ames strain. To locate SNPs with the highest specificity for the Ames strain, we screened these SNPs across a panel of isolates that were identified as close genetic relatives of the Ames strain by using high-resolution MLVA subtyping (6, 17; M. N. Van Ert and P. Keim, unpublished data). Here we demonstrate the use of strain-specific canSNPs for subtyping *B. anthracis*, which can assist in epidemiological and forensic investigations.

**MATERIALS AND METHODS**

**Strains used in this study.** We selected a panel of 89 *B. anthracis* strains that represent 89 unique MLVA genotypes as described by Keim et al. (6) and includes an Ames strain isolate (genotype 62). In addition, we included two isolates that were previously identified as Ames (Δ*Ames* [A0814; *Ames* cured of pX01 and pX02]), a clinical isolate from the 2001 letter attack ([A2012]), and an additional 10 isolates previously shown to be genetically related to the Ames strain by sequencing to screen the 32 chromosomal SNPs across a panel of DNA from isolates that were genetically similar to the Ames strain and from a distantly related *B. anthracis* strain (Pasteur [14]) plasmids. The authors indicated that the SNPs exhibited specificity for the Ames strain in relation to the other strains involved in the study, including the closely related A0394 “Texas goat” isolate (see Table S1 in the supplemental material). Based upon these results, two SNPs located on the pX01 plasmid (PS-1 and PS-32) and one SNP located on the pX02 plasmid (PS-52) were selected as candidates.

**Initial screening of Ames strain-specific chromosomal SNPs.** We used DNA sequencing to screen the 32 chromosomal SNPs across a panel of DNA from the Ames strain, isolates that were genetically similar to the Ames strain and from a distantly related *B. anthracis* strain (Vollum; see Table S1 in the supplemental material). DNA sequencing was performed according to the manufacturer’s protocol (Applied Biosystems, Foster City, CA). The SNPs located on the pX01 and pX02 plasmids were screened against these strains using allelic discrimination real-time PCR.

**TaqMan MGB allelic discrimination assays.** TaqMan minor-groove binding (MGB) allelic discrimination assays were designed around 4 of the 32 chromosomal SNPs which exhibited the highest specificity for the Ames strain. In addition, TaqMan MGB allelic discrimination assays were designed around three plasmid SNPs, PS-1, PS-32, and PS-52, reported by Read et al. (17). TaqMan MGB probes and primers for the four chromosomal SNPs and the pX01 and pX02 plasmids were designed by Applied Biosystems. Probe and primer sequences are listed in Table 1. Each 10.0-µl reaction mixture contained 1X ABI Universal Master Mix, 250 nM of each probe, 600 nM each of forward and reverse primers, and 1.0 µl of approximately 350 pg/µl template DNA. For all assays, thermal cycling parameters were 50°C for 2 min and 95°C for 10 min, followed by 40 to 50 cycles of 95°C for 15 s and 60°C for 1 min.

**TABLE 1. Sequences of primers and TaqMan MGB probes**

<table>
<thead>
<tr>
<th>SNP</th>
<th>Primers 5' to 3' sequences for</th>
<th>TaqMan MGB probe&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SNP change and genome position</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS-1</td>
<td>F-TGATGGTTTTGTATTCTTATG</td>
<td>FAM-AAGGGTCΔCAAATC</td>
<td>G⇒A</td>
<td>NC_003980</td>
</tr>
<tr>
<td></td>
<td>R-CACTTTGGTGTTGATGTTAAG</td>
<td>VIC-AAGGGTCΔCAAATC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PS-52</td>
<td>F-GTATCCTGAAAATATAAAGTGTA</td>
<td>VIC-ATTAAGGACTCCTTTTGT</td>
<td>A⇒C</td>
<td>NC_003981</td>
</tr>
<tr>
<td></td>
<td>AATTGAAAAATGGAGTA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R-GATCTTCAACGGAATATACCTATCA</td>
<td>FAM-AAGGACTCCTΔTTTTGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTAATTATACATATATATATATATATAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branch1-7</td>
<td>F-TCACTCTCAATGACATCGCA</td>
<td>FAM-CAAAAACAATAΔCCCCTTT</td>
<td>C⇒A</td>
<td>NC_003997</td>
</tr>
<tr>
<td></td>
<td>R-TTGTITGGAAAGCAGGATAACTTT7</td>
<td>VIC-CAAAACAAATAΔCCCCTTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TATG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branch1-26</td>
<td>F-GACGGAGCAACACAGAGA</td>
<td>FAM-ATAGCTTTTTTTTTTATTC</td>
<td>T⇒C</td>
<td>NC_003997</td>
</tr>
<tr>
<td></td>
<td>R-CGGTGTGAAATGAGCAGTTGAAAA</td>
<td>VIC-ATAGCTTTTTTTATTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TTTC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branch1-28</td>
<td>F-AATATCTTTCTCATACAAAGGGCGAC</td>
<td>FAM-CTGGTGTAGTTTTTAC</td>
<td>T⇒G</td>
<td>NC_003997</td>
</tr>
<tr>
<td></td>
<td>TACT</td>
<td>VIC-CTGGTGTAGTTTTTAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R-CCATAATGTCTGTCTCAAAATC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branch1-31</td>
<td>F-GAAGGGAAACGGGAAAAGCT</td>
<td>VIC-CGTTTCACATGGCAT</td>
<td>A⇒G</td>
<td>NC_003997</td>
</tr>
<tr>
<td></td>
<td>ACCT</td>
<td>VIC-CGTTTCACATGGCAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R-GAGTTTTCTAATCAGTTGAAAAAGTG</td>
<td>FAM-TGGTTTCACATAGCAT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Note that all probes have a minor-groove binder, nonfluorescent quencher attached to the 3' end. FAM, 6-carboxyfluorescein. The Ames probes were labeled with FAM except in the case of PS-1, which was labeled with VIC. Underlined letters indicate SNP location.

<sup>b</sup> SNP changes are expressed as “Ames ⇒ non-Ames.”
TABLE 2. Profiles of chromosomal and plasmid-based single-nucleotide polymorphisms<sup>a,b</sup>

| SNP   | A0462 | A1115 | A1117 | A0394 | A0728 | A0584 | A0576 | A0585 | A0580 | A0720 | A0724 | A0729 | A0488
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PS-52</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-7</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-26</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-28</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Branch1-31</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Branch1-8</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>Branch1-16</td>
<td>B</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-18</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Branch1-22</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-23</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Branch1-27</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Branch1-30</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Branch1-1</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-6</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Branch1-10</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Branch1-19</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-20</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Branch1-21</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-24</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Branch1-25</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-29</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Branch1-31</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-2</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-3</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-4</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>PS-32</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Branch1-17</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-19</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-20</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>Branch1-21</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-25</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Branch1-29</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-31</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Branch1-2</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-3</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Branch1-4</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

<sup>a</sup> Underlined letters indicate alleles shared with the Ames strain. One of the chromosomal SNPs was removed from the analysis since our sequence data indicated that the SNP was not specific for the phylogenetic branch containing the Ames strain (data not shown).

<sup>b</sup> SNP profiles for the chromosomal SNPs Branch1-1 to Branch1-31 were determined by DNA sequencing using flanking primers, whereas SNP profiles for the plasmid-based SNPs were determined by TaqMan MGB genotyping.

RESULTS

Identification and screening of Ames strain-specific SNPs. Sequence analysis of the 32 chromosomal SNPs located along the B. anthracis Ames strain lineage (see Fig. S1 in the supplemental material; 15) indicated four SNPs (Branch1-7, Branch1-26, Branch1-28, and Branch1-31) separated the Ames MLVA type from all of the remaining genetic relatives (Table 2). Notably, this includes the separation of Ames from three closely related Texas isolates. These four SNPs were selected for further testing based upon their specificity for the Ames strain. Of the three TaqMan MGB assays designed around plasmid SNPs, the PS-52 SNP assay separated the Ames MLVA type from all of the remaining genetic relatives, whereas the PS-1 SNP assay exhibited less specificity and grouped the Ames strain with a single Chinese strain and the three closely related Texas isolates (Table 2). The PS-32 SNP assay exhibited much lower specificity for the Ames strain (Table 2).

TaqMan allelic discrimination genotyping assays. To confirm the specificity of the Ames SNP assays (for SNPs, Branch1-7, Branch1-26, Branch1-28, Branch1-31, PS-1, and PS-52) for the Ames strain, we used TaqMan dual-probe allelic discrimination to genotype 88 globally and genetically diverse B. anthracis isolates. The cycle threshold ($C_T$) values observed in the real-time data were similar among the assays designed around the chromosomal and pX02 SNPs (for Branch1-7, the mean was 21.1 $C_T$, the standard deviation [SD] was ±0.96, and the range was 18.3 to 23.3; for Branch1-26, the mean was 23.6 $C_T$, the SD was ±0.91, and the range was 20.5 to 26.0; for Branch1-28 the mean was 21.5 $C_T$, the SD was ±0.93, and the range was 18.8 to 23.2; for Branch1-31 the mean was 22.6 $C_T$, the SD was ±1.01, and the range was 19.5 to 24.7; and for PS-52, the mean was 22.9 $C_T$, the SD was ±1.2, and the range was 19.6 to 27.7). The assay designed around the pX01 SNP PS-1 exhibited lower $C_T$ values (the mean was 18.8 $C_T$, the SD was ±0.90, and the range was 16.2 to 21.3).
The post-PCR allelic discrimination plot from these experiments illustrates the unambiguous separation of genotypes for the Ames strain and remaining “non-Ames” diverse 88 strains. All six SNP assays separated the Ames strains from the remaining 88 globally diverse \textit{B. anthracis} genotypes (Fig. 1A to F). Although the PS-1 SNP assay separates Ames from the remaining 88 global strains, it will cluster Ames with three closely related Texas isolates and a single Chinese isolate (TaqMan data not shown; Table 2 presents SNP profile data).

\textbf{Level of detection of TaqMan MGB genotyping assays.} The TaqMan MGB allelic discrimination assays were used to reliably detect and genotype samples containing as little as 100 fg of template DNA (ca. 17 genome equivalents), whereas samples containing 10 fg of template DNA (ca. 1.7 genome equivalents) could be detected and genotyped sporadically. Fig. 2A illustrates the real-time amplification plots generated from analysis of a 10-fold serial dilution of template DNA from the Ames strain with the Branch1-7 SNP TaqMan assay. The post-PCR allelic discrimination plot from the same experiment (Fig.
2B) illustrates the unambiguous separation of genotypes for both the Ames strain and the non-Ames (Vollum) strain over a range of template DNA quantities from 100 pg to 10 fg. C_T values for all of the assays run across this serial dilution of template DNA, and numerous replicates of 10 fg input DNA are reported in Tables S2 and S3 in the supplemental material.

DISCUSSION

The identification of strain-specific SNP markers in *B. anthracis* permits the development of rapid diagnostics to greatly assist in the investigation of biocrimes and natural outbreaks. Forensic and epidemiological investigations can require the analysis of hundreds or even thousands of specimens, including environmental and clinical samples. During the 2001 anthrax outbreak, the CDC used MLVA to subtype over 100 *B. anthracis* isolates and several dozen clinical samples (3). Significantly, these researchers used MLVA to include or exclude *B. anthracis* strains from the epidemiological investigation. Since MLVA requires PCR, post-PCR processing, electrophoresis, and fragment sizing, there is a practical limit to the number of samples that can be strain typed in a timely manner. In contrast, Ames-specific real-time assays allow thousands of DNA samples per day to be included or excluded as the Ames strain.

The combination of Ames strain-specific SNPs and real-time PCR allows for the development of assays that are both specific and sensitive. Five SNPs we identified were highly specific and were observed in only the Ames strains (A2012, A0462, and A0814 [chromosome only]), whereas a single SNP located on the pX01 plasmid (PS-1) partitioned the Ames strain with three closely related Texas isolates and a single Chinese strain. The pXO1 SNP assay is still valuable because of its ability to monitor the plasmid composition of an isolate. Our data suggest that the SNPs which define the Ames genetic lineage are evolutionarily stable since SNPs found in the Ames strain, but not close genetic relatives, were strain specific when evaluated across a more comprehensive isolate collection (i.e., the diverse 88 *B. anthracis* strains).

The stability of these SNPs as diagnostic markers is likely a function of the low mutation rates of nucleotide substitutions in *B. anthracis* (10^{-10} per site per generation [21]), the recent evolutionary derivation of *B. anthracis* from *Bacillus cereus*, and the lack of recombination due to the highly clonal nature of *B. anthracis* (8, 15). The rarity of these mutational events within *B. anthracis* limits the likelihood of a SNP mutating again to a novel or ancestral state and lowers the probability of observing a false positive in any one of the SNP assays (i.e., a
non-Ames strain sharing the allele with Ames). The lack of recombination ensures that the few mutational events are restricted to a particular clade. The combinatorial use of the six Ames SNP assays vastly increases confidence in the inclusion/exclusion of a sample as the Ames strain since the likelihood that a different lineage of \textit{B. anthracis} would conversely evolve to match this six-locus genotype is vanishingly small (e.g., the product of the six individual locus mutation rates is about \(\sim 10^{-60}\)).

The Ames strain-specific PCR assays were sensitive, allowing reliable detection and genotyping of samples containing 100 fg of template DNA, which we estimate corresponds to approximately 17 genome equivalents. We routinely use one chromosomal SNP assay (Branch1-31) and two plasmid SNP assays (PS-1 and PS-52) on blinded proficiency samples and have achieved 100% accuracy in discriminating Ames versus non-Ames strain templates over a range of DNA levels (>260 samples [data not shown]). It is important to point out that the distinct genotypic clustering of samples in the endpoint analysis and the similar \(C_T\) values observed among \textit{B. anthracis} strains argue against the presence of unexpected sequence variants (unexpected polymorphisms in probe/primer sites) in our diverse \textit{B. anthracis} strain panel. If present, unexpected sequence variation among diverse strains can dramatically impact assay performance that ultimately results in less sensitivity and higher thresholds of detection.

Examining SNPs with various specificity for the Ames strain also provides insights on the phylogeny of the Ames evolutionary lineage. Based upon our SNP analyses, the nearest relatives to the Ames strain include two isolates cultivated from cattle in Texas in 2001 (Alex Hoffmaster, CDC) and a previously reported isolate cultivated from a goat in Texas in 1997 (17). These findings are supported by MLVA data and are consistent with the knowledge that the Ames strain was originally isolated from cattle in Jim Hogg County, Texas, in 1981 (Martin Hugh-Jones, personal communication). The isolation of two of these Ames relatives in 1997 and 2001 suggests the Ames lineage is endemic in this region. It is interesting that the remaining genetic near relatives to the Ames strain were isolated from human outbreaks in China (1957 and 1962), suggesting that the Chinese and Texas isolates share a common ancestral lineage. International trade of spore-contaminated products (bone meal, wool, etc.) can facilitate such long-distance dispersal of \textit{B. anthracis} and may explain this geographically distant relationship.

Evaluating SNP data in relation to higher resolution genotyping data (MLVA genotypes) provides for fine-scale phylogenetic hypotheses that can be used to guide strain-specific SNP discovery. The applicability of this approach is not limited to \textit{B. anthracis}, and we expect that this strategy will be used to identify strain-specific SNPs in other pathogens and biowarfare agents. For example, high-resolution variable-number tandem repeat markers have been identified in \textit{Yersinia pestis} (9) and whole-genome sequence comparisons of this pathogen are currently underway. As whole-genome sequences and high-resolution molecular typing methods become available for biological threat agents, this strategy could assist efforts in the forensic and biodefense arena to develop assays coupling low-level detection with highly specific strain identification.

In conclusion, our strategy permits the identification of SNPs that are diagnostic for narrowly defined \textit{B. anthracis} genetic lineages. The Ames strain-specific SNP assays are based on six selected genomic differences between Ames and its genetic relatives and work at the single-molecule limit. Hence, we are approaching the theoretical limits for both specificity (single nucleotide) and sensitivity (single molecule) for strain subtyping. The theory of strain identification must be coupled with the practical limitation of laboratory analysis, necessitating extensive validation and quality control (7). With this caveat, the use of strain-specific SNPs for low-level, high-throughput genotyping may prove to be a valuable tool to assist in real-world forensic and public health studies.

**ACKNOWLEDGMENTS**

This work was supported by funding from the United States Department of Homeland Security and by the Federal Bureau of Investigation.

We thank Alex Hoffmaster and Bijiang Wang for providing Ames-related strains and Marty Johnson and Lilly Wong (Applied Biosystems) for their technical assistance in assay design.

**REFERENCES**


