Advances in Bioscience and Biotechnology, 2014, 5, 669-677
Published Online June 2014 in SciRes. http://www.scirp.org/journal/abb
http://dx.doi.org/10.4236/abb.2014.57079

Application of High-Throughput Sequencing: Discovery of Informative SNPs to Subtype *Bacillus anthracis*

Guillaume Girault*, Simon Thierry, Emeline Cherchame, Sylviane Derzelle

University Paris-Est, ANSES, Animal Health Laboratory, Bacterial Zoonoses Unit, Maisons-Alfort, France
Email: *guillaume.girault@anses.fr, simon.thierry@anses.fr, emeline.cherchame@anses.fr, sylviane.derzelle@anses.fr

Received 17 April 2014; revised 29 May 2014; accepted 17 June 2014

Abstract

Single Nucleotide Polymorphisms (SNPs) are the most common and abundant genetic variation found in the genome of any living species, from bacteria to humans. In bacterial genotyping, these evolutionarily stable point mutations represent important DNA markers that can be used to elucidate deep phylogenetic relationships among worldwide strains, but also to discriminate closely related strains. With the advent of next generation sequencing (NGS) technologies, affordable solutions are now available to get access to the complete genome sequence of an organism. Sequencing efforts of an increasing number of strains provide an unprecedented opportunity to create comprehensive species phylogenies. In this study, a comparative analysis of 161 genomes of *Bacillus anthracis* has being conducted to discover new informative SNPs that further resolves *B. anthracis* SNP-based phylogenetic tree. Nine previously unpublished SNPs that define major groups or sub-groups within the *B. anthracis* species were selected and developed into SNP discriminative assays. We report here a cost-effective high-resolution melting-based genotyping method for the screening of 27 canonical SNPs that includes these new diagnostic markers. The present assays are useful to rapidly assign an isolate to one sub-lineages or sub-groups and determine its phylogenetic placement on the *B. anthracis* substructure population.

Keywords

SNPs, *Bacillus anthracis*, Genotyping, HRM, Phylogeny

1. Introduction

Bacillus anthracis is a zoonotic pathogen that primarily affects grazing herbivores, but can be used as an agent of

*Corresponding author.

How to cite this paper: Girault, G., Thierry, S., Cherchame, E. and Derzelle, S. (2014) Application of High-Throughput Sequencing: Discovery of Informative SNPs to Subtype *Bacillus anthracis*. *Advances in Bioscience and Biotechnology*, **5**, 669-677. http://dx.doi.org/10.4236/abb.2014.57079

bioterrorism. This Gram-positive endospore-forming bacterium is the causative agent of anthrax, an infectious and life threatening diseases in all mammals, including humans. *B. anthracis* evolved as a highly fit clonal pathogen that has spread throughout the world and became ecologically established as local, genetically distinct populations [1].

In both microbiological forensics and epidemiological investigations, obtaining reliable information regarding the identification and source of a suspicious strain is often essential to set up effective responses, as highlighted by the Amerithrax investigations following the so-called 2001 "anthrax letter attacks" in the United States [2]. The principal genomic markers used for genotyping a highly monomorphic bacterium such as *B. anthracis* are variable number tandem repeats (VNTRs) [3]-[5], single nucleotide repeats (SNRs) [6]-[8] and SNPs [9]. Since 2007, the genotyping of a set of 13 canonical SNPs has emerged as the gold standard method to be used for determining phylogenetic relationships within this bacterium [10]-[13]. SNPs are generated by nucleotide substitutions, probably associated to errors during DNA replication that are not repaired. These rare events occur at an estimated rate of approximately 10^{-10} changes per nucleotide per generation in the *B. anthracis* species [14]. Because of their low mutation rates, SNPs represent ideal targets for phylogenetic analyses. SNPs are evolutionarily stable markers that are unlikely to mutate back to their ancestral state [9]. This stability can be used for defining canonical markers representative of groups of related strains, such as phylogenetic sub-groups [9].

The substructure of the *B. anthracis* population is divided into three major clades (A, B and C), with further subdivisions into 13 major lineages or groups [10] [12] [13]. The A-radiation has known the most dramatic dispersal and clonal expansion. It represents nowadays the majority of anthrax cases reported around the world [3] [15] [16]. The eight A-lineages or groups exhibit distinct geographical patterns across the five continents. Most basal of the A-clade is the African A.Br.005/006 group, consistent with an "out of Africa" hypothesis for *B. anthracis*. Soon after this divergence, the Vollum-clade appeared, with isolates currently ecologically established in Iran, Pakistan and Afghanistan, for instance. The highly successful TransEurasian (TEA) group (A.Br.008/009) is slightly more recent and spread across Europe and Asia from which a derived lineage (A.Br. WNA) was introduced into North America. Another recent bifurcation led to the A.Br.003/004, A.Br. Australia-94 and Ames/Sterne (A.Br.001/002) groups which contain isolates found, respectively, in South America (A.Br.003/004), Asia and part of Europe (A.Br.Aust94 and A.Br.001/002) [1] [10]. Collections from western and south-central Asia (Turkey, India) and western China are dominated by genotypes belonging to the A.Br.Aust94 lineage. Further into central and eastern China the genotypes are dominated by isolates belonging to the A.Br.001/002 group and its derived A.Br.Ames lineage. The B clade is divided into two genetically distinct lineages with geographical restricted repartition. The B.Br.001/002 group (including the derived B.Br.Kruger lineage) is ecologically established in Southern Africa, in particular in the Kruger National Park (South Africa), where it co-exists with other strains from the A-clade [10] [17]. The B.Br.CNEVA lineage is exclusively found in continental Europe [10] [18]-[21]. The C-clade is composed of an uncommon lineage, C.Br.A1055, clustering only three collection's isolates of unknown origin [10] [22].

Whole genome sequencing is the ultimate genotyping tool, giving the highest possible resolution, with comparison information spanning from distant phylogenetic relationships to the highest level of subtyping. Identification of thousands of SNPs retrieved from compiled NGS sequences facilitates high-resolution strain tracking and provides the level of discrimination required for microbial forensics. Canonical SNP typing and whole genome comparisons of multiple strains will soon become the method of choice in *B. anthracis* genotyping [23]. We have recently reported the high throughput sequencing and phylogeographical analysis of 122 strains of *B. anthracis* isolated in France [24]. The resulting whole-genome SNPs analysis determined relationships existing between French strains at a level of discrimination never reached before and reconstructed the French population substructure in details. Eight new canonical SNPs have been identified that allow classifying French *B. anthracis* isolates at a higher resolution [24]. In this study, we compared the chromosomal sequences of 161 globally diverse strains, focusing on identifying additional informative SNPs, to provide further resolution into current SNP-based typing of *B. anthracis*. Twenty-seven SNP discrimination assays, including nine newly described markers, are reported. The PCR-HRM-based method presented here is a cost-effective tool to genotype the highly clonal *B. anthracis* species.

2. Materials and Methods

2.1. Extraction of Whole-Genome SNPs among *B. anthracis* Strains

In this study, the genome of one African A.Br.005/006-affiliated strain (IEMVT 89) and a hundreds of French

strains (n = 123) belonging to 3 groups or lineages (A.Br.001/002, A.Br.011/009 (A.Br.008/009 sub-group), and B.Br.CNEVA) [24] were whole-genome compared to thirty-seven globally diverse genomes publically available in database [25] (**Table 1**). Ames Ancestor [GenBank: AE017334.2] was used as reference genome. Sequences alignment and whole-genome SNPs discovery were carried out using the BioNumerics version 7.0 software (Applied Maths, Belgium) and its Chromosome Comparisons module. SNPs were next filtered to avoid erroneous and biased data. Missing sequences, ribosomal operons, VNTR loci and contiguous SNPs in a 10 pb window were eliminated from the analysis.

Table 1. Whole genome sequences available in public database used in this study.

Strain	Country	canSNP	Accession number
Ames Ancestor	USA	A.Br.Ames	NC_007530.2
A2012	USA	A.Br.Ames	AAAC00000000.1
Ames	USA	A.Br.Ames	NC_003997.3
A0248	USA	A.Br.Ames	NC_012659.1
Sterne	South Africa	A.Br.001/002	AE017225.1
A0389	Indonesia	A.Br.001/002	ABLB00000000.1
Ba103	Japan	A.Br.001/002	DRR000183 (SRA)
A16	China	A.Br.001/002	CP001970.1
V770-Np1-R (ATCC 14185)	Israël	A.Br.003/004	AZQO00000000.1
CZC5	Zambie	A.Br.005/006	BAVT00000000.1
Tsiankovskii	Soviet Union	A.Br.008/011	ABDN00000000.2
Ba 3154	Bulgaria	A.Br.008/011	ANFF00000000.1
Ba 3166	Bulgaria	A.Br.008/011	ANFG00000000.1
Heroin Ba4599	Scotland	A.Br.008/011	AGQP00000000.1
UR-1	Germany	A.Br.008/011	ALNY00000000.1
Carbosap	Italy	A.Br.011/009	ANAO00000000.1
Sen2Col2	Africa	A.Br.011/009	CAVC000000000.1
Sen3	Africa	A.Br.011/009	CAVD000000000.1
Gmb1	Africa	A.Br.011/009	CAVE000000000.1
Australia 94	Australia	A.Br.Australia94	AAES00000000.1
9080 G	Georgia	A.Br.Australia94	AZUE00000000.1
52 G	Georgia	A.Br.Australia94	AZUF00000000.1
8903 G	Georgia	A.Br.Australia94	AZUD00000000.1
CDC684	USA	A.Br.Vollum	NC_012581.1
H9401	Korea	A.Br.Vollum	NC_017729.1
A0488	UK	A.Br.Vollum	ABJC00000000.1
Vollum	UK	A.Br.Vollum	AAEP00000000.1
USA6153	USA	A.Br.WNA	AAER00000000.1
A0174	Canada	A.Br WNA	ABLT00000000.1
A0193	USA	A.Br WNA	ABKF00000000.1
A0442	South Africa	B.Br.001/002	ABKG00000000.1
SVA11	Sweden	B.Br.001/002	CP006742.1
BF1	Germany	B.Br.CNEVA	AMDT00000000.1
CNEVA-9066	France	B.Br.CNEVA	AAEN00000000.1
A0465	France	B.Br.CNEVA	ABLH00000000.1
Kruger B	South Africa	B.Br.KrugerB	AAEQ00000000.1
A1055	USA	C.Br.A1055	AAEO00000000.1

2.2. Phylogenetic Analysis

A minimum spanning tree was drawn in BioNumerics by using the filtered whole genome sequencing SNP data as input. Nodes were automatically numbered by the software. Canonical SNPs that defining major group, sub-groups or clusters of strains along the SNP tree were identified by searching for key signatures with allelic states shared only by these selected subgroups.

2.3. SNP Discrimination Assays by HRM

After identification, primer sequences suitable for HRM analysis were designed for each SNP marker using the Primer 3+ software (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/) (**Table 2**). Short amplicons (<100 pb) were selected. High-resolution melting (HRM) is a post-PCR technique that determines with high precision the melt profile of PCR products. A new generation dye is incorporated into a double-stranded DNA. Using a slow constant increase in temperature, fluorescence acquisition allows the distinction between two different populations of amplicons.

Amplification was performed on the ViiA7™ Real-Time PCR System (Life Technologies) using the LightCycler® 480 High Resolution Melting Master Mix (Roche Diagnostics). The reaction mixture consisted of 0.2 µM of each primer, 1 × LightCycler® 480 HRM master mix and 2.5 mM $MgCl_2$ in a 10 µl final volume. The following parameters were used: 10 min at 95˚C were followed by 40 cycles consisting of 10 s at 95˚C, 10 s at 58˚C and 20 s at 72˚C. Samples were next heated to 95˚C for 30 s, cooled down to 65˚C for 1 min and heated from 65˚C to 88˚C at a rate of 1˚C/s with 25 acquisitions/˚C. HRM data were analyzed by the ViiA7™ Software (version 1.2.1).

3. Results

3.1. *In Silico* SNP Typing

Thirty-seven *B. anthracis* genomes available in public NCBI database or Sequence Read Archive (SRA) were used for this study [25] (**Table 1**). These globally diverse genomes were first subjected to *in silico* canSNP typing using the 14 canSNPs references set previously published [10] [12] to determine their phylogenetic placement on the *B. anthracis* canSNP tree (**Table 1**). French strains were previously shown to belong to A.Br.011/009 (a sub-group of the TEA population) (n = 32), A.Br.001/002 (n = 24) and B.Br.CNEVA (n = 67) lineages or groups. The African IEMVT 89 strain is affiliated to A.Br.005/006 [24].

3.2. Phylogenetic Analysis

A total of 161 genomes of *B. anthracis* were aligned to the Ames ancestor reference sequence and whole-genome compared with a focus on SNPs discovery. SNPs that define the major A.Br.Aust94 lineage (four genomes), A.Br.003/004 group (one genome) and A.Br.005/006 group (two genomes) were identified and one canonical marker selected for each subpopulation (**Table 2**, **Figure 1**). Interestingly, the whole-genome SNP analysis further resolved the A.Br.001/002 subpopulation (32 genomes) into two distinct phylogenetic sub-groups which were named A01 and A02 (**Figure 1**, inset A). The A01 sub-group (also termed "Ames sub-group") radiates very shortly after the A01-A02 divergence (1 SNP at position 515111) into at least 5 sub-branches. The first one contains the A.Br.Ames lineage (strains Ames Ancestor, Ames, A2012 and A0248). The second one is composed by the A0389 strain. The third one includes the Japanese strain Ba103 and one old French strain (CIP A211). The two last branches are composed of, respectively, two old French isolates (CIP 81.89 and CIP 53.169), and the A16 Chinese strain. The A02 sub-group (also termed "Sterne sub-group") clusters the Sterne vaccine strain and all recent strains isolated in France from the Doubs department, including the 08-8_20 genome [24]. Noteworthy to mention, the Carbosap vaccine was found to be phylogenetically related to five French genomes positioned within the third branch of the six sub-branches previously discovered within the A.Br.011/009 sub-groups. SNPs specific to these A.Br.011/009 branches were also selected (**Table 2**, **Figure 1**, inset B).

3.3. canSNP Discrimination Assays Using HRM

Twenty-seven canSNP discrimination assays, including nine unpublished markers, were designed using the

Table 2. canSNPs and primer sequences used for HRM analysis.

canSNP name	Target lineage	Position*	Forward Primer (5' - 3')	Reverse Primer (5' - 3')	SNP	Reference
A.Br.001	A.Br.Ames	182106	CAAGCGGAACCAAATTTAATCTTT	TTCACCGTACGTCATTGTATAATACG	T/C	[10]
A.Br.002	A.Br.Ames, A.Br.001/002	947760	AACGATACCTAAAATCGATAAAG	GGCAGAAGGAGCAAGTAATGTT	G/A	[10]
A01	A.Br.Ames, A.Br.01	515111	CGTGGCGTAAAAATAAAGATCC	TCGTTTGCTGCAACATAATTTC	A/G	[11]; this study
A02	A.Br.02	240050	AGCTGCTGAAAATGATGTAGCA	GAACCATCCCAAAGTTTGATTG	T/C	[11]; this study
Ba20	A.Br.02 (Fr-Doubs)	3434997	AGCGAGCCAATTTTGGAACCGA	AGGCGGGATTGTTGGTGGATGT	A/C	[24]
A.Br.003	A.Br.Ames, A.Br.001/002, A.Br.Aust94	1493157	GCTACTGTCATTGTATAAAAACCTCCTTT	CGCTTGCCAAGCTTTTTTTC	A/G	[10]
A94	A.Br.Aust94	569295	CCCTGCACCTAACATAACAACA	CCGCAATGTACCCACTGTATAA	C/T	[11]; this study
A.Br.004	A.Br.Ames, A.Br.001/002, A.Br.Aust94, A.Br.003/004	3600659	CCGATACCAGTAAACGACGACAT	CTGGAATTGGTGGAGCTATGGA	T/C	[10]
A03	A.Br.003/004	3956945	CGAAAGAACATATTGCACTCGTAA	GCGTACAAGTACAGGCTCACCT	A/G	this study
A05	A.Br.005/006	405303	CTATATTTTGGACTTGTGCGACAG	GCAATGTTTTGGGAGCTTAGTG	C/T	this study
A.Br.006	A-Clade	162509	CCGGAAATTGCTATTAGAACGAA	TCCCAATCTAGCGTTTTTAAGTTCA	C/A	[10]
A.Br.007	A.Br.Vollum	266439	TTGGTAACGAGACGATAAACTGAATAA	GCCTTGGATTGGCGATTG	T/C	[10]
A.Br.008	A.Br.008/009	3947248	TTCGCAACTACGCTATACGTTTTAGAT	CAAACGGTGAAAAAGTTACAAATATACG	T/G	[10]
A.Br.009	A.Br.WNA	2589823	GGCAATCGGCCACTGTTT	GGGTTTCTACTGTGTATGTTGTTAATAAAAAG	A/G	[10]
A.Br.011	A.Br.011/009	2552486	CTAAAAAGAAACGAATTCCCGCTGA	CGATAAAAATCGGAATTGAAGCAGGAG	G/A	[12]; [13]
A11/1 (Ba21)	A.Br.011/009 (Branch 1 Fr)	3562427	AGCAAAAAGTCGGCAAAGAA	ACAGAGCTTCCTCCGAACTG	A/C	[24]
A11/2 (Ba22)	A.Br.011/009 (Branch 2 Fr)	820195	AGTGGTGCAATCCCAATTTC	CGCAGCAATATTCGCTATCA	T/C	[24]
A11/3	A.Br.011/009 (Branch 3)	1975689	CACATTGTCAATATTAAATCGCTGA	TGCTTCTCCACCTAAATCAAATG	G/A	this study
A11/4	A.Br.011/009 (Branch 4)	2571863	TCCTACACCTGTTTCACCACAA	AAGCTTGTATCACCAACTGATGC	C/T	this study
A11/5	A.Br.011/009 (Branch 5)	5021456	TATTTGTTGAAAGAGCGTCATCC	TGTGTCAGTCTCCGCATACAATA	T/C	this study
A11/6	A.Br.011/009 (Branch 6)	1391599	ACCAAGTAATTAAGTTCCAAGCTATTG	ATTCGCCTTACAAAATGGTTCTC	C/T	this study
B.Br.001	B.Br.KrugerB	1455279	TGCATGCTTCTTCTTACAGAGTAGTTAAT	CGGTCATAAAAGAAATCGGTACAA	T/C	[10]
B.Br.002	B.Br.KrugerB, B.Br.001/002	1056740	TGTTGCACCTTCTGTGTTCGTT	GTAGTGGCTTCACCGAATGGA	G/T	[10]
B.Br.003	B-Clade	1494269	CATTTATTCGCATAGAAGCAGATGA	TGTGCCATCAAATAACTCTTTTCTCAA	G/A	[10]
B.Br.004	B.Br.CNEVA	69952	GAAGTTAAGTATCAACCAGCAGAAGAAA	CCGCCGCCTTGAGCTT	T/C	[10]
Ba 19	B.Br.CNEVA (France)	2573536	CATATATTTTCACCTCTTTTATGAACA	GATAAAAGGCTGTCGGATGG	A/G	[24]
A/B.Br.001	C.Br.A1055	3697886	GAAGGTCTCCAATTTGGATTTAAAAT	CGTGTGAACCTTTCGGTAAATAGTC	A/G	[10]

*Position based on the Ames Ancestor genome (NC_007530.2).

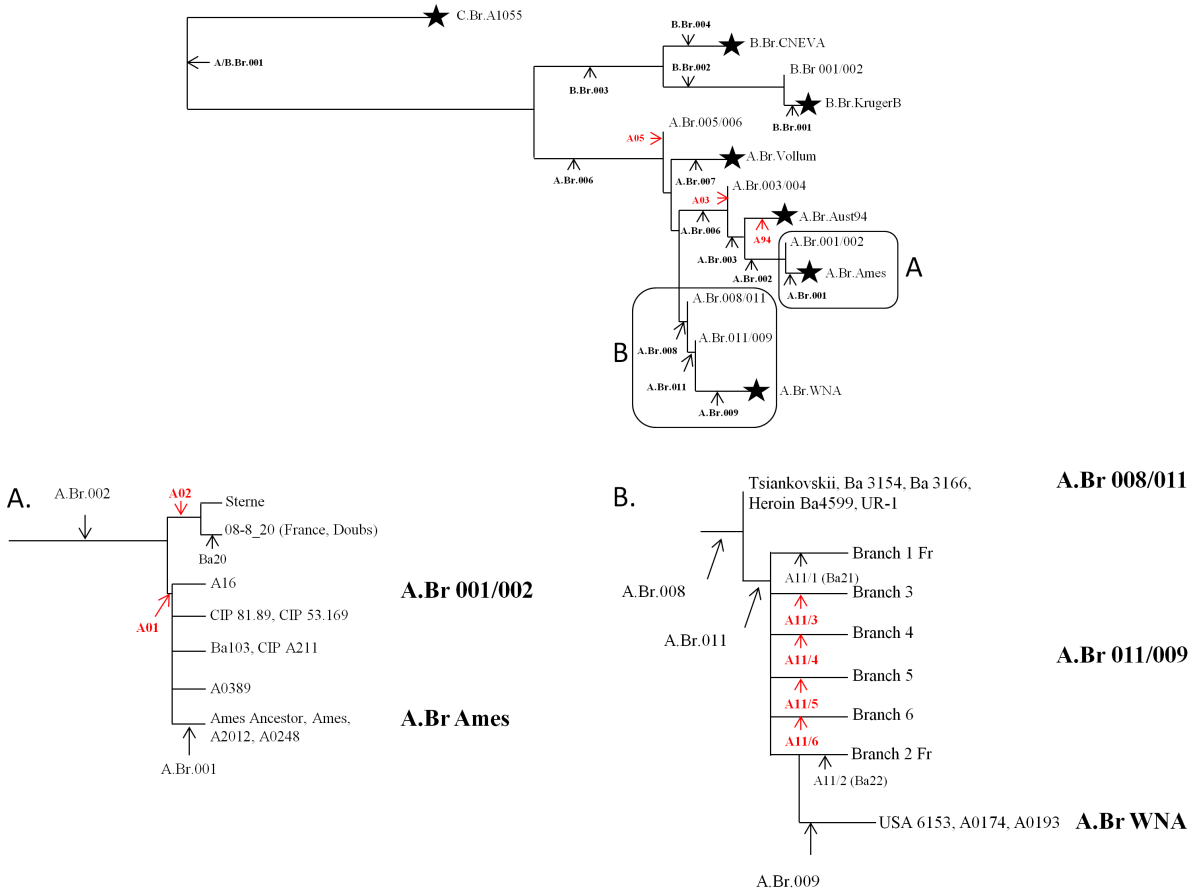

Figure 1. Phylogeny of the major canSNP groups of *B. anthracis*. Dendrogram of 14 canSNP analysis of *Bacillus anthracis* isolates after Van Ert *et al*. (2007) and Marston *et al*. (2011). Canonical SNP positions and names that define lineages, groups or sub-groups are indicated in black arrows (previously published canSNP) and red arrows (newly discovered canSNPs). The A.Br.001/002 group is now subdivided into two sub-groups: A.Br.01 and A.Br.02 (inset A). The A.Br.011/009 sub-group is further resolved into six branches as previously described for the French TEA subpopulation (inset B). These new groups are named after the canSNP marker and assay that define their position.

HRM-PCR technology (**Table 2**). Novel assays targeted four canSNPs specific to the A.Br.011/009 sub-branches 3 to 6 (A11/3, A11/4, A11/5, A11/6), two canSNPs specific to the new A01 and A02 sub-groups within A.Br.001/002 (A01, A02) and three canSNPs specific to the main A.Br.Aust94, A.Br.003/004 and A.Br.005/006 subdivisions (A94, A03 and A05, respectively). These diagnostic assays were successfully validated against all DNA samples of our *B. anthracis* collection (about 250), including a hundred of non-French *B. anthracis* DNAs of diverse origins (described by 10 canSNP typing). Average melting temperatures calculated for each bi-allelic assay are reported in **Table 3**. All non-sequenced genomes fell on the correct group (data not shown), confirming HRM-based assays' specificity.

4. Discussion

Cost is an important issue for all diagnostic assays. In this respect, real-time PCR platforms and high-resolution DNA melting analysis (HRM) are attractive technologies for SNP-interrogation. HRM techniques can determine with high precision the melt profile of PCR products using accurate fluorescence data acquisition over small temperature increments [26]. Two standard primers are used to amplify short segments flanking each SNP. The melting profile of the resulting amplicon is characteristic of its GC content, *i.e*., a substitution of a G or C to an A or T reduces the melting temperature (*Tm*) while a substitution of an A or T to a G or C increases the *Tm*. Based on its simplicity, low cost, non-destructive nature, high sensitivity and specificity, the popularity of HRM analysis has grown considerably in the last few years. HRM analysis is an effective alternative to other

Table 3. Melting temperature of new canonical SNPs.

canSNP name	Target	Allele	Tm values (˚C)
A01	A.Br.Ames, A.Br.001/002 subgroup A01	G	77.2 ± 0.1
	Others	A	76.6 ± 0.1
A02	A.Br.001/002 subgroup A02	C	75.9 ± 0.1
	Others	T	75.3 ± 0.1
Ba20	A.Br.001/002 subgroup A02 (France, Doubs)	C	80.4 ± 0.1
	Others	A	79.8 ± 0.1
A94	A.Br.Aust94	T	77.3 ± 0.1
	Others	C	78.0 ± 0.1
A03	A.Br.003/004	G	76.6 ± 0.1
	Others	A	75.8 ± 0.1
A05	A.Br.005/006	T	76.4 ± 0.1
	Others	C	77.6 ± 0.1
A11/1 (Ba21)	French A.Br.011/009 Branch 1	C	77.9 ± 0.1
	Others	A	77.1 ± 0.1
A11/2 (Ba22)	French A.Br.011/009 Branch 2	C	77.5 ± 0.1
	Others	T	76.9 ± 0.1
A11/3	A.Br.011/009 Branch 3	A	72.1 ± 0.1
	Others	G	72.9 ± 0.1
A11/4	A.Br.011/009 Branch 4	A	74.6 ± 0.1
	Others	G	75.8 ± 0.1
A11/5	A.Br.011/009 Branch 5	C	76.2 ± 0.1
	Others	T	75.3 ± 0.1
A11/6	A.Br.011/009 Branch 6	T	75.8 ± 0.1
	Others	C	76.6 ± 0.1
Ba19	French B.Br.CNEVA group	G	75.8 ± 0.1
	Others	A	75.2 ± 0.1

PCR-based methods such as Dual Probe TaqMan PCR assays [27] and Mismatch Amplification Mutation Assays (MAMA) [28] [29]. HRM is an ideal format for scoring a small number of SNPs. PCR-HRM assays can be rapidly designed around SNPs to determine the extent to which each marker varies in the *B. anthracis* population.

Over the last decade, significant research efforts have been undertaken to develop genotyping methods with increased resolving power for *B. anthracis* strain differentiation. The absence of horizontal transfer of genetic material, a slow rate of accumulation of mutations and the paucity of large genome rearrangements within the genome of *B. anthracis* [9] make SNPs markers of choice for genotyping a species. The advent of next-generation sequencing technologies and comparative genome analyses offers a powerful approach for identifying rare polymorphisms for the unbiased typing of highly clonal pathogens. Since 2007, and the identification by Van Ert *et al*. of 13 "canonical" SNPs representing key phylogenetic positions along the *B. anthracis* phylogenetic SNP tree, a growing number of diagnostic SNPs has been discovered and used to survey *B. anthracis* diversity

in nature [11] [24] [30]-[32]. In this study, a same but large-scale approach, comparing 161 *B. anthracis* genomes of diverse origins, were conducted to discover novel informative SNPs that can further discriminate within lineages or groups of strains. Nine novel DNA signatures were selected and developed into SNP discrimination assays. Combined with previously established canSNPs [10]-[12] [24], the set of 27 markers described here allows a more precise and reliable phylogenetically assignment of any *B. anthracis* strain into 20 sub-lineages or sub-groups. The present method will contribute to improve our ability to characterize *B. anthracis* strains and to react rapidly when the identity and origin of a suspicious strain need to be established.

Acknowledgements

This work was supported by the French National Research Agency (ANR) and Direction Générale de l'Armement (DGA) (project number ANR11ASTR0007). GG is a PhD student co-supported by ANSES and DGA grants.

This work has benefited from the facilities and expertise of the high throughput sequencing platform of IMAGIF (Centre de Recherche de Gif, http://www.imagif.cnrs.fr).

References

[1] Keim, P., Okinaka, R. and Pearson, T. (2013) The Evolution and Global Dissemination of *Bacillus anthracis*. *The International Conference on Bacillus anthracis, B. cereus, and B. thuringiensis*, Victoria, 1-5 September 2013.

[2] Rasko, D.A., Worsham, P.L., Abshire, T.G., Stanley, S.T., Bannan, J.D., Wilson, M.R., Langham, R.J., Decker, R.S., Jiang, L., Read, T.D., *et al.* (2011) *Bacillus anthracis* Comparative Genome Analysis in Support of the Amerithrax Investigation. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 5027-5032. http://dx.doi.org/10.1073/pnas.1016657108

[3] Keim, P., Price, L.B., Klevytska, A.M., Smith, K.L., Shupp, J.M., Okinaka, R., Jackson, P.J. and Hugh-Jones, M.E. (2000) Multiple-Locus Variable-Number Tandem Repeat Analysis Reveals Genetic Relationships within *Bacillus anthracis*. *Journal of Bacteriology*, **182**, 2928-2936. http://dx.doi.org/10.1128/JB.182.10.2928-2936.2000

[4] Lista, F., Faggioni, G., Valjevac, S., Ciammaruconi, A., Vaissaire, J., le Doujet, C., Gorge, O., De Santis, R., Carattoli, A., Ciervo, A., *et al.* (2006) Genotyping of *Bacillus anthracis* Strains Based on Automated Capillary 25-Loci Multiple Locus Variable-Number Tandem Repeats Analysis. *BMC Microbiology*, **6**, 1-15.

[5] Le Flèche, P., Hauck, Y., Onteniente, L., Prieur, A., Denoeud, F., Ramisse, V., Sylvestre, P., Benson, G., Ramisse, F. and Vergnaud, G. (2001) A Tandem Repeats Database for Bacterial Genomes: Application to the Genotyping of *Yersinia pestis* and *Bacillus anthracis*. *BMC Microbiology*, **1**, 2.

[6] Stratilo, C.W., Lewis, C.T., Bryden, L., Mulvey, M.R. and Bader, D. (2006) Single-Nucleotide Repeat Analysis for Subtyping *Bacillus anthracis* Isolates. *Journal of Clinical Microbiology*, **44**, 777-782. http://dx.doi.org/10.1128/JCM.44.3.777-782.2006

[7] Kenefic, L.J., Beaudry, J., Trim, C., Daly, R., Parmar, R., Zanecki, S., Huynh, L., Van Ert, M.N., Wagner, D.M., Graham, T., *et al.* (2008) High Resolution Genotyping of *Bacillus anthracis* Outbreak Strains Using Four Highly Mutable Single Nucleotide Repeat Markers. *Letters in Applied Microbiology*, **46**, 600-603. http://dx.doi.org/10.1111/j.1472-765X.2008.02353.x

[8] Garofolo, G., Ciammaruconi, A., Fasanella, A., Scasciamacchia, S., Adone, R., Pittiglio, V. and Lista, F. (2010) SNR Analysis: Molecular Investigation of an Anthrax Epidemic. *BMC Veterinary Research*, **6**, 11. http://dx.doi.org/10.1186/1746-6148-6-11

[9] Keim, P., Van Ert, M.N., Pearson, T., Vogler, A.J., Huynh, L.Y. and Wagner, D.M. (2004) Anthrax Molecular Epidemiology and Forensics: Using the Appropriate Marker for Different Evolutionary Scales. *Infection, Genetics and Evolution*: *Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, **4**, 205-213.

[10] Van Ert, M.N., Easterday, W.R., Huynh, L.Y., Okinaka, R.T., Hugh-Jones, M.E., Ravel, J., Zanecki, S.R., Pearson, T., Simonson, T.S., U'Ren, J.M., *et al.* (2007) Global Genetic Population Structure of *Bacillus anthracis*. *PLoS One*, 2, Article ID: e461.

[11] Kuroda, M., Serizawa, M., Okutani, A., Sekizuka, T., Banno, S. and Inoue, S. (2010) Genome-Wide Single Nucleotide Polymorphism Typing Method for Identification of *Bacillus anthracis* Species and Strains among *B. cereus* Group Species. *Journal of Clinical Microbiology*, **48**, 2821-2829. http://dx.doi.org/10.1128/JCM.00137-10

[12] Marston, C.K., Allen, C.A., Beaudry, J., Price, E.P., Wolken, S.R., Pearson, T., Keim, P. and Hoffmaster, A.R. (2011) Molecular Epidemiology of Anthrax Cases Associated with Recreational Use of Animal Hides and Yarn in the United States. *PLos One*, **6**, Article ID: e28274. http://dx.doi.org/10.1371/journal.pone.0028274

[13] Price, E.P., Seymour, M.L., Sarovich, D.S., Latham, J., Wolken, S.R., Mason, J., Vincent, G., Drees, K.P., Beckstrom-Sternberg, S.M., Phillippy, A.M., *et al*. (2012) Molecular Epidemiologic Investigation of an Anthrax Outbreak among Heroin Users, Europe. *Emerging Infectious Diseases*, **18**, 1307-1313. http://dx.doi.org/10.3201/eid1808.111343

[14] Vogler, A.J., Busch, J.D., Percy-Fine, S., Tipton-Hunton, C., Smith, K.L. and Keim, P. (2012) Molecular Analysis of Rifampicin Resistance in *Bacillus anthracis* and *Bacillus cereus*. *Antimicrobial Agents and Chemotherapy*, **46**, 511-513. http://dx.doi.org/10.1128/AAC.46.2.511-513.2002

[15] Keim, P., Kalif, A., Shupp, J., Hill, K., Travis, S.E., Richmond, K., Adair, D.M., Hugh-Jones, M., Kuske, C.R. and Jackson, P. (1997) Molecular Evolution and Diversity in *Bacillus anthracis* as Detected by Amplified Fragment Length Polymorphism Markers. *Journal of Bacteriology*, **179**, 818-824.

[16] Keim, P., Gruendike, J.M., Klevytska, A.M., Schupp, J.M., Challacombe, J. and Okinaka, R. (2009) The Genome and Variation of *Bacillus anthracis*. *Molecular Aspects of Medicine*, **30**, 397-405. http://dx.doi.org/10.1016/j.mam.2009.08.005

[17] Smith, K.L., DeVos, V., Bryden, H., Price, L.B., Hugh-Jones, M.E. and Keim, P. (2000) *Bacillus anthracis* Diversity in Kruger National Park. *Journal of Clinical Microbiology*, **38**, 3780-3784.

[18] Derzelle, S., Laroche, S., Le Fleche, P., Hauck, Y., Thierry, S., Vergnaud, G. and Madani, N. (2011) Characterization of Genetic Diversity of *Bacillus anthracis* in France by Using High-Resolution Melting Assays and Multilocus Variable-Number Tandem-Repeat Analysis. *Journal of Clinical Microbiology*, **49**, 4286-4292. http://dx.doi.org/10.1128/JCM.05439-11

[19] Garofolo, G., Serrecchia, L., Corro, M. and Fasanella, A. (2011) Anthrax Phylogenetic Structure in Northern Italy. *BMC Research Notes*, **4**, 273. http://dx.doi.org/10.1186/1756-0500-4-273

[20] Pilo, P., Perreten, V. and Frey, J. (2008) Molecular Epidemiology of *Bacillus anthracis*: Determining the Correct Origin. *Applied and Environmental Microbiology*, **74**, 2928-2931. http://dx.doi.org/10.1128/AEM.02574-07

[21] Gierczyński, R., Kałuzewski, S., Rakin, A., Jagielski, M., Zasada, A., Jakubczak, A., Borkowska-Opacka, B. and Rastawicki, W. (2004) Intriguing Diversity of *Bacillus anthracis* in Eastern Poland—The Molecular Echoes of the Past Outbreaks. *FEMS Microbiology Letters*, **239**, 235-240.

[22] Sue, D., Marston, C.K., Hoffmaster, A.R. and Wilkins, P.P. (2007) Genetic Diversity in a *Bacillus anthracis* Historical Collection (1954 to 1988). *Journal of Clinical Microbiology*, **45**, 1777-1782.

[23] Derzelle, S. and Thierry, S. (2013) Genetic Diversity of *Bacillus anthracis* in Europe: Genotyping Methods in Forensic and Epidemiologic Investigations. *Biosecurity and Bioterrorism*: *Biodefense Strategy*, *Practice*, *and Science*, **11**, S166-S176.

[24] Girault, G., Blouin, Y., Vergnaud, G. and Derzelle, S. (2014) High-Throughput Sequencing of *Bacillus anthracis* in France: Investigating Genome Diversity and Population Structure Using Whole-Genome SNP Discovery. *BMC Genomics*, **15**, 288.

[25] NCBI. http://www.ncbi.nlm.nih.gov/

[26] Erali, M., Voelkerding, K.V. and Wittwer, C.T. (2008) High Resolution Melting Applications for Clinical Laboratory Medicine. *Experimental and Molecular Pathology*, **85**, 50-58.

[27] Hurtle, W., Bode, E., Kulesh, D.A., Kaplan, R.S., Garrison, J., Bridge, D., House, M., Frye, M.S., Loveless, B. and Nordwood, D. (2004) Detection of the *Bacillus anthracis gyrA* Gene by Using a Minor Groove Binder Probe. *Journal of Clinical Microbiology*, **42**, 179-185.

[28] Easterday, W.R., Van Ert, M.N., Zanecki, S. and Keim, P. (2005) Specific Detection of *Bacillus anthracis* Using a TaqMan Mismatch Amplification Mutation Assay. *Biotechniques*, **38**, 731-735.

[29] Birdsell, D.N., Pearson, T., Price, E.P., Hornstra, H.M., Nera, R.D., Stone, N., Gruendike, J., Kaufman, E.L., Pettus, A.H., Hurbon, A.N., *et al.* (2012) Melt Analysis of Mismatch Amplification Mutation Assays (Melt-MAMA): A Functional Study of a Cost-Effective SNP Genotyping Assay in Bacterial Models. *PLoS One*, **7**, Article ID: e32866.

[30] Kenefic, L.J., Pearson, T., Okinaka, R.T., Schupp, J.M., Wagner, D.M., Hoffmaster, A.R., Trim, C.B., Chung, W.K., Beaudry, J.A., Jiang, L., *et al.* (2009) Pre-Columbian Origins for North American Anthrax. *PLoS One*, **4**, Article ID: e4813.

[31] Simonson, T.S., Okinaka, R.T., Wang, B., Easterday, W.R., Huynh, L., U'Ren, J.M., Dukerich, M., Zanecki, S.R., Kenefic, L.J., Beaudry, J., *et al.* (2009) *Bacillus anthracis* in China and Its Relationship to Worldwide Lineages. *BMC Microbiology*, **9**, 71.

[32] Okinaka, R.T., Henrie, M., Hill, K.K., Lowery, K.S., Van Ert, M., Pearson, T., Shupp, J., Kenefic, L., Beaudry, J., Hofstadler, S.A., *et al.* (2008) Single Nucleotide Polymorphism Typing of *Bacillus anthracis* from Sverdlovsk Tissue. *Emerging Infectious Diseases*, **14**, 653-656.

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.