academicJournals

Vol. 15(45), pp. 2538-2546, 9 November, 2016
DOI: 10.5897/AJB2016.14902
Article Number: 0A64E6961539
ISSN 1684-5315
Copyright © 2016
Author(s) retain the copyright of this article
http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Pan-genome analysis of Senegalese and Gambian strains of *Bacillus anthracis*

M. Mbengue[1*], F. T. Lo[1], A. A. Diallo[1], Y. S. Ndiaye[1], M. Diouf[1] and M. Ndiaye[2]

[1]National Laboratory for Research on Animal Diseases (LNERV – ISRA) - Hann – Dakar – P. O Box 2057, Senegal.
[2]Biocellular laboratoty for Research on Microbiology and Rickettssiology, Faculty for Sciences and Technology - Dakar University (UCAD) – Senegal.

Received 4 August, 2015; Accepted 10 October, 2016

Bacillus anthracis **is the causative agent of anthrax, and it is classified as "category A" biological weapon. There were six available complete genomes (A0248, Ames, Ames Ancestor, CDC684, H0491 and Sterne). Here, one Gambian and two Senegalese strains (Gmb1, Sen2Col2 and Sen3) were added. In this work, the pan-genome of *B. anthracis* was studied based on nine strains and using bioinformatics tools as Cluster of Orthologous Groups (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG). Thereafter, *B. anthracis* pan-genome having 2893 core genes and 85 accessory genes was estimated. With Mauve method, the pan-genome of *B. anthracis* was verified and it was found to be very narrow and clonal. To have confidence in this study, different tools were used to compare and validate the results. All of the tools yielded the same results; the addition of the Senegalese and Gambian strains did not change the nature of the *B. anthracis* pan-genome (2893 core genes and 85 accessory genes), which had a core/pan-genome ratio of 99%. The closed nature of the pan-genome of *B. anthracis* (the core genome) represents 99% of the pan-genome size. The hypothesis that *B. anthracis* had a closed pan-genome was hereby validated.**

Key words: *Bacillus anthracis,* Senegalese, Gambian strains, pan-genome.

INTRODUCTION

Anthrax was the first disease to be attributed to a specific microbe, thanks to Davaine in 1863 (Scarlata et al., 2010) and the first animal infection for which we had a vaccine by Pasteur in 1881 (Scarlata et al., 2010). In 1876, Koch discovered for the first time a bacterium which has the capacity to transform into spores: *Bacillus anthracis* (Scarlata et al., 2010). *B. anthracis*, in the Firmicutes phylum and belonging to the *Bacillus cereus* group (Kuroda et al., 2010), is a Gram positive spore-forming bacterium (Wang et al., 2012), which is able to survive in extreme and unfriendly environmental conditions as high levels of radiation or extreme temperature (Wang et al., 2012) and can stay viable in the soil for a long time (Sweeney et al., 2011). *B. anthracis* is the causative agent of anthrax (Kuroda et al., 2010), a zoonosis. Cattle and horses are mainly sensitive (Scarlata et al., 2010).

*Corresponding author. E-mail: mbenguem@yahoo.fr.

Author(s) agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Humans can be infected by various routes: ingestion, inhalation of spores or through the skin (Kuroda et al., 2010). There are four clinical syndromes for anthrax disease (Sweeney et al., 2011): cutaneous anthrax (95% of the reported cases), gastrointestinal anthrax (due to contaminated food), inhalational anthrax, and injectional anthrax. *B. anthracis* is classified as a "Category A" potential biothreat (Wang et al., 2012). Indeed, due to the stability of its spores, the high level pathogenicity and lethality and its capacity to be infected by the inhalational route (Rasko et al., 2011), this bacterium represents a bioterrorism weapon. In these days, one bioterrorist attack was done in 2001 in the United States (Scarlata et al., 2010) using a strain of *B. anthracis*, the potential source of which was identified based on genomic analysis. Earlier, they had an attack in USSR in 1979, with an anthrax epidemic through an atmospheric contamination from a military laboratory (Guillemin, 2002; Scarlata et al., 2010). The 2001 event led to an increase of the research about *B. anthracis* and anthrax (Imperiale and Casadevall, 2011) and allows the emergence of new detection system (Wang et al., 2012).

The first genome sequencing study of multiple stains was published in 2005 on *Streptococcus agalactiae* (Tettelin et al., 2005) and, since then, such pangenomic studies have increased quickly. On working on pan-genomes, allowed a comparison between different species or strains, and it is defined like the pool of all the genes present in all the studied genomes. This can be divided into different parts: the core genome (genes present in all the genomes), accessory genes which are present in some genomes and unique genes (genes present only in one of the studied genomes). A pan-genome can be closed or open, depending of the capacity of the species to acquire new genes (Tettelin et al., 2005) and of the age of the initial clone.

Senegalese and Gambian strains of *B. anthracis* have not been compared to the other strains (Read et al., 2002). In this study, analysis of the *B. anthracis* pan-genome was carried out based on three African strains (two from Senegal and one from Gambia) and on six reference genomes [Ames (Read et al., 2003), Ames Ancestor (Ravel et al., 2009), A0248, CDC684 (Okinaka et al., 2011), H9401 (Chun et al., 2012), and Sterne]. The present study shows that African strains were very closely related to the other strains and presented a closed pan-genome, as already shown in previous studies.

MATERIALS AND METHODS

Bacteriological studies

Cells from various organs were cultured in a liquid medium consisting of ordinary broth. After seeding, the medium was incubated at 37°C for 24 h. The isolation ensues on sheep blood agar (blood culture) which was incubated as earlier stated. Gram stain is performed from isolates, as well as the study of biochemical characteristics.

Sequencing

The sequencing strategies of the three strains *B. anthracis* Sen2col2, Sen3 and Gmb1 were carried out through the SOLiD 4_Life technologies in New Generation Sequencing (NGS) technologies (Figure 2). Sequencing of the Sen2col2, Sen3 and Gmb1 strains of *B. anthracis* were performed using the SOLiD 4_Life Technology's New Generation Sequencing technology. The paired end library was constructed from 1 lg of purified genomic DNA from each strain. The sequencing was carried out to 50935 base pairs (bp) using SOLiDTM V4 chemistry on one full slide associated with 96 other projects on an Applied Biosystems SOLiD 4 machine (Applied Biosystems, Foster City, CA, USA). All 96 genomic DNA samples were barcoded with the module 1 to 96 barcodes provided by Life Technologies (Paisley, UK). The libraries were pooled in equimolar ratios, and emPCR (PCR by emulsion) was performed according to the manufacturer's specifications, using templated bead preparation kits on the EZ bead automated Emulsifier, Amplifier and Enricher E80 system for full-scale coverage. A total of 708 million P2-positive beads were loaded onto the flow cell for the run and the output read length was 85 bp, as expected (50935 bp). The three *B. anthracis* genomes (Sen2col, Sen3 and Gmb1) were sequenced through 3.2^{E} + 6, 3.1^{E} + 6 and 3.9^{E} + 6 barcode reading which led to 273, 262, and 382 Mb of data, respectively. The global sequencing of these three genomes resulted in 917 Mb of data.

Basic genomic data

The complete genomic sequences of the six references strains are available on the NCBI: Ames (NC_003997.3), Ames Ancestor (NC_007530.2), A0248 (NC_012659.1), CDC684 (NC_012581.1), H9401 (NC_017729.1), and Sterne (NC_005945.1). Our strains of interest came from Senegal (Sen2Col2 and Sen3) and from Gambia (Gmb1). They were isolated in 2010. The first one (Sen2Col2) was isolated from lungs of a 6 years old ostrich. The second one (Sen3) came from lungs, liver, spleen and blood of a Touabire race sheep. The last one, Gmb1, was isolated on trypanotolerant zebu cattle's blood (Table 1). The sequences of these three Senegalese and Gambian strains (Sen2Col2, Sen3 and Gmb1) were obtained in reference to SOLiD data.

Genomic analysis

Cluster of Orthologous Groups (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG)

CAMERA (Sun et al., 2011) is a bioinformatics portal where several kind of analysis can be done. It was used to generate the COG data. COG (Tatusov et al., 2001) is a common tool, used to assign functional annotations to proteins. These proteins were classified into categories (the list is available at http://www.ncbi.nlm.nih.gov/COG/old/palox.cgi?fun=all). To get KEGG (Ogata et al., 1999) data and to investigate metabolic pathways, the KAAS (Moriya et al., 2007) (KEGG automatic annotation server) online tool was used. In KEGG, the proteins were classified into classes and subclasses.

Alignments, pan-genome

First of all, two kinds of alignments were performed: a global genome alignment with MAUVE (Darling et al., 2010). With MAUVE (Figure 1) and its backbone output file (Sheppard et al., 2013), the proportion of core genome depending on the pan-genome size was calculated to evaluate the close or open nature of the pan-genome.

Table 1. Characteristics of *Bacillus anthracis strains* (samples).

Code	Strains	Country	Host	Samples origin
G192	Sen2Col2	Senegal	Ostrich 6 Years old	Lungs
G193	Sen3	Senegal	Mouton de race Touabire	Lungs, blood
G194	Gmb1	Gambia	Bovin de race Ndama Trypanotolérante	Blood
Ref strains	Ames	USA	Vache morte	/
Ref strains	Ames Ancestor	USA	Beefmaster female 14 years old	/
/	A0248	USA	Human	/
/	H9401	Korea	Human	/
/	CDC684	USA	Human	/
/	Sterne	UK	/	/

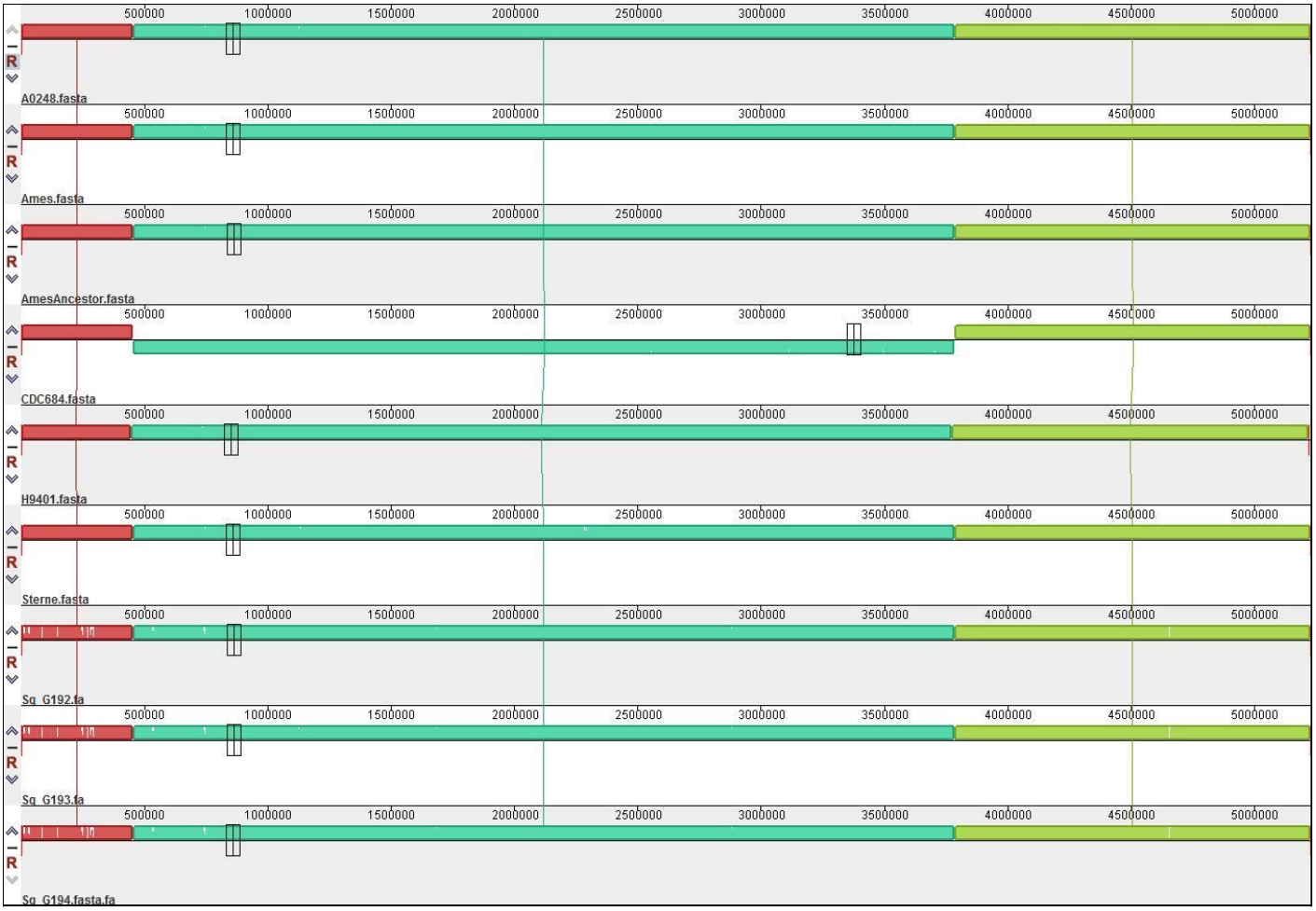

Figure 1. MAUVE global alignement.

Then, OrthoMCL (Chen et al., 2006) was used to obtain a list of orthologs to determine the pan-genome composition (core, accessory and unique genes). MeV (Saeed et al., 2006, 2003) (Multi Experiment Viewer) was used to best visualize the accessory genes distribution and to perform a hierarchical clustering (Figure 7). Clustering of the strains was based on the distribution of all the Cluster of Orthologous Groups categories: J, translation, ribosomal structure and biogenesis; K, transcription; L, replication, recombination and repair; B, chromatin structure and dynamics; D, cell cycle control, cell division, chromosome partitioning; O, post-translational modification, protein turnover, chaperones; M, cell wall/membrane/envelope biogenesis; N, cell motility and secretion; P, inorganic ion transport and metabolism; T, signal transduction mechanisms. SNPs contained in the core genome were also worked on.

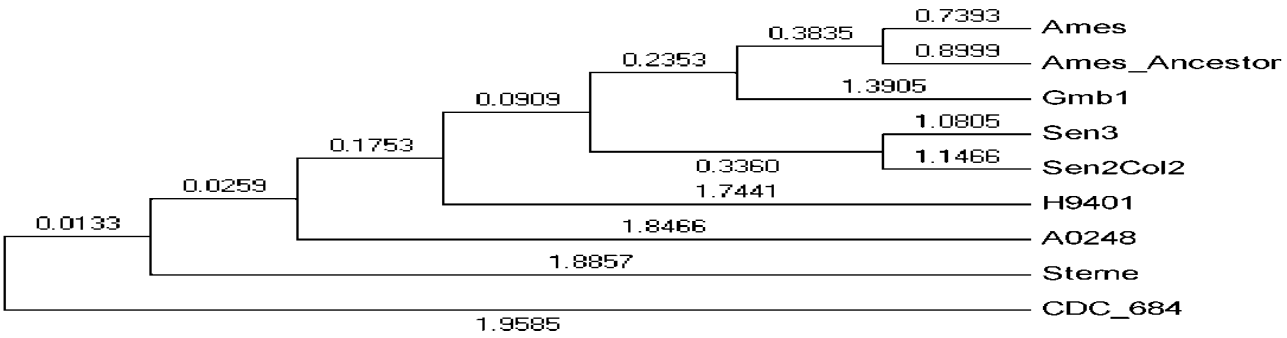

Figure 2. Phylogeny based on complete genome.

Table 2. General characteristics for strains.

Strain	Plasmid	Size (Mb)	GC (%)	Proteins
B*acillus anthracis* str. Ames	/	5.23	35.4	5,328
Bacillus anthracis str. CDC 684	pX01, pX02	5.23	35.4	5,579
Bacillus anthracis str. Sterne	/	5.23	35.4	5,289
Bacillus anthracis str. 'Ames Ancestor'	pX01, pX02	5.23	35.4	5,208
Bacillus anthracis str. A0248	pX01, pX02	5.23	35.4	5,041
Bacillus anthracis str. H9401	BAP1, BAP2	5.22	35.4	5,479
Bacillus anthracis str.Sen2Col2 (G192)	pX01,pX02	5.23	35,2	5,487
Bacillus anthracis str.Sen3 (G193)	pX01,pX02	5.23	35,2	5,493
Bacillus anthracis str.Gmb1 (G194)	pX01,pX02	5.23	35,3	5,502
Bacillus anthracis str.Sen1 (G195)	/	/	/	/

Therefore, we get back the sequences of all the core genes (based on the OrthoMCL part), thanks to a Perl script and used SNPs finder (Song et al., 2005) for the core genome tree.

RESULTS

Culture of *B. anthracis*

On ordinary broth after 24 h of incubation at 37°C, the appearance of flakes at the bottom of the tube was observed, leaving a supernatant clear enough. Mobile bacilli in long chains Gram-positive was not observed after examination. The pathogenicity test on Balb/C mouse was confirmed after 6 h strains inoculation; all of them were dead. Cultural, morphological and biochemical characteristics were studied in detail using conventional methods.

Pan-genome analysis: The obtention of genomic sequences results and their bio informatic analysis has allowed knowing the structure for pan-genome

The pan-genome is composed of 2893 core genes, 7 unique genes, and 85 accessory genes (Figure 6). First, we looked at unique genes. Five in Sterne (2 not found, one conserved hypothetical protein, EmrB/QacA family drug resistance transporter and zinc-binding dehydrogenase), 1 in CDC 684 (not found on the NCBI) and 1 in H9401 (yfeT DNA-binding transcriptional regulator) were found. Then, we looked in details on the 85 accessory genes (Figure 6). The three African strains and CDC684 possessed almost all the accessory genes, whereas A0248 owned only 20 accessory genes out of 85. The half of the accessory genes was annotated as hypothetical proteins (Figure 6). It was noticed that Ames Ancestor owned 42 accessory genes, whereas, its non-virulent version, Ames, owned more (59). The hierarchical clustering (Figure 6) showed again the same two groups found, thanks to COG (Table 3, Figures 3 and 5) and KEGG (Figure 4) (one with Ames, Ames ancestor and A0248, the second with all the other strains). Moreover, the core/pan-genome ratio was done and core genome represented 99% of the pan-genome (Table 2), was found showing again the high rate of conservation between the nine strains. Finally, the SNPs at core genome level were studied. We found 896 SNPs, that is, 32% of the total number of SNPs (2786 SNPs found in comparing all the genomes); and a transition/transversion bias of 0.32. The very small rate of SNPs, the low transition/transversion bias and the very high proportion of the core genome function as the pan-genome (Table

Table 3. COG distribution between various *B. anthracis* strains.

Categories	Ames ancestor	A0248	Ames	Sterne	CDC684	H9401	Sen2Col2	Sen3	Gmb1	Class description
J	209	208	210	211	211	210	210	211	212	Translation ribosomal Structure
K	381	376	383	393	391	387	391	391	391	Transcription
L	155	156	156	159	160	160	159	162	162	Replication and transcription
B	1	1	1	1	1	1	1	1	1	Chromatin structure and dynamics
D	40	40	40	44	44	44	44	45	45	Cell cycle control, cell division
O	101	100	101	103	102	103	102	102	102	Posttranslational modification, protection
M	206	203	215	246	241	241	246	247	251	Cell wall/membrane/ envelope biogenesis
N	48	47	50	60	59	59	224	61	61	Cell motility and secretion
P	226	222	224	246	246	246	54	245	245	Defense mechanism
T	203	202	204	225	224	223	106	224	224	Signal translation mechanism
U	47	46	46	52	54	54	54	55	54	Intracellular traficknng
V	93	93	94	107	106	105	106	106	107	Defense mechanisms
W	0	0	0	0	1	1	0	0	0	Extracellular structure
Z	1	1	1	1	1	1	1	1	1	Cytoskeleton
C	198	198	198	206	205	207	205	205	207	Energy production and conversion
G	248	247	249	268	268	267	270	270	270	Carbohydrate transport and metabolism
E	377	373	375	407	405	403	409	410	409	Amino acide transport and metabolism
F	121	120	122	130	130	131	131	131	131	Nucleotide transport and metabolism
H	169	168	169	173	175	175	175	175	175	Coenzyme transport and metabolism
I	123	121	123	128	129	126	129	126	130	Lipid transport and metabolism
Q	72	72	73	78	82	81	82	81	83	Secondary metabolites biosynthesis
R	543	541	552	587	582	580	582	580	588	General function prediction only
S	427	420	431	442	444	439	444	439	468	Function unknown

4) showed that *B. anthracis* is an ancient protein (probably very older than 150 years). It was believed that the lack of gene transfer and defense mechanisms (CRISPRs) observed in intracellular bacteria suggests that *B. anthracis* multiplies only as a pathogen and that its life in soil is exclusively dormant.

DISCUSSION

In comparing the three African strains to the others, it was noticed that all the Senegalese and Gambian strains are closer from CDC684. In this work, a validation of that was given as previously shown (Tettelin et al., 2005); the pan-genome of *B. anthracis* is narrow. To have confidence in our study, different tools were used in order to compare and validate the results. All go in the same sense: the addition of the Senegalese and Gambian strains did not change the closed nature of this pan-genome (2893 core genes and 85 accessory genes), with a core/pan-genome ratio of 99%. This core/pan-genome ratio is very close from the other human clonal pathogens (Table 1) as *Rickettsia rickettsii.* However, there is discordance between the presence of a mobilome; which is a structure localized in the pXO1 and pXO2 plasmids and contained five transposases, one phage and no CRISPRs and the fact to have a closed pan-genome. Nevertheless, *B. anthracis* derived from *B. cereus* group, a sympatric species that is not intracellular. Therefore, *B. anthracis* may become allopatric (Table 4). This can be explained by the fact that *B. anthracis* is an ancient bacterium (at least 150

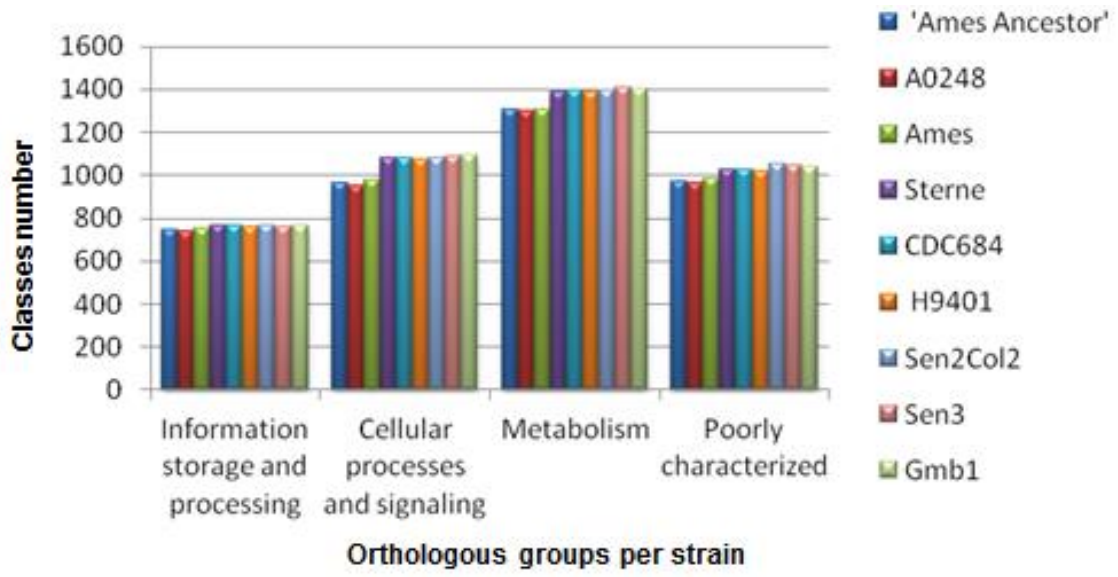

Figure 3. Cluster of Orthologous Groups (COG) classes distribution between *B. anthracis* strains.

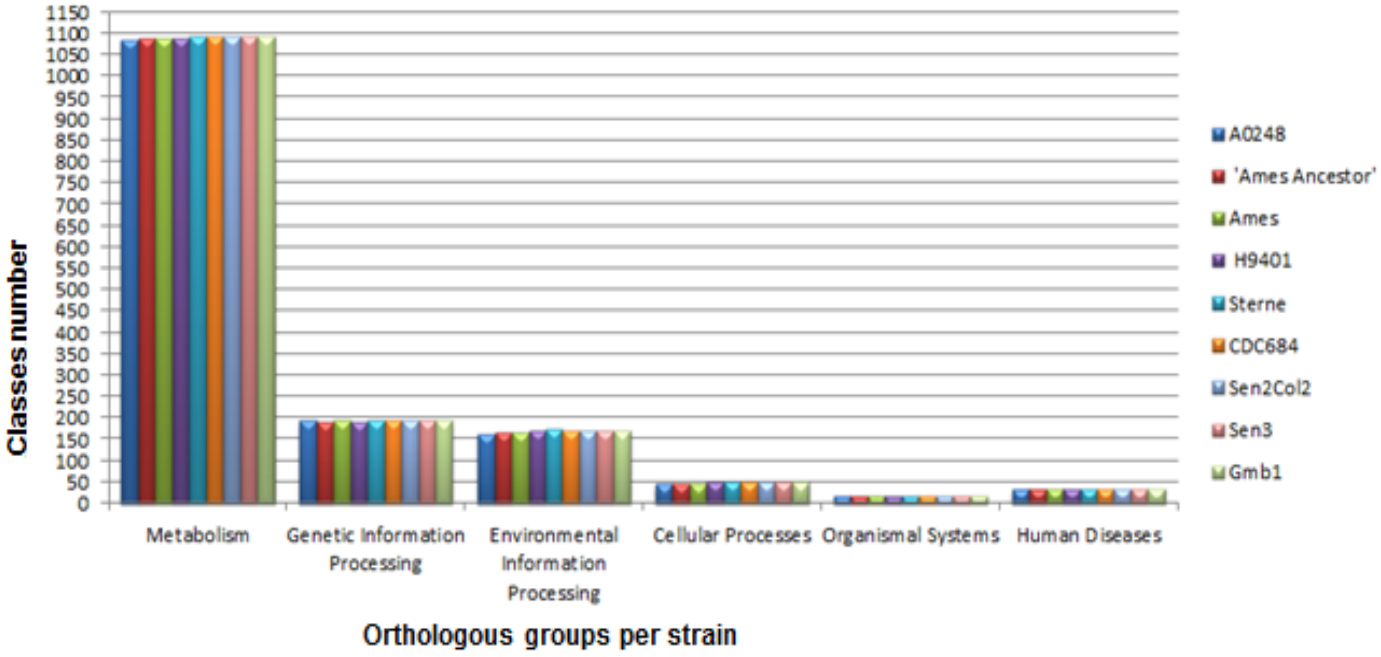

Figure 4. Distribution of KEEG *B. anthracis* classes.

years) which evolve. This hypothesis was tested in studying SNPs based on the core genome content. Only 2786 SNPs in total with 896 in the core genome were found. Moreover, the transition/transversion process is very small (0.32). This lack of SNPs may validate our hypothesis of the evolution of this species. *B. anthracis* is an ancient clone which is stabilized with the time and which present a conserved pan-genome.

Conclusion

B. anthracis was discovered 150 years ago, but kept the same genomic content. We are in a case of a very closed pan-genome with species which do not live in the environment. Due to the lengthy spore phase of its life cycle, *B. anthracis* evolved very slowly and has a very narrow pan-genome, despite its apparent soil ecological

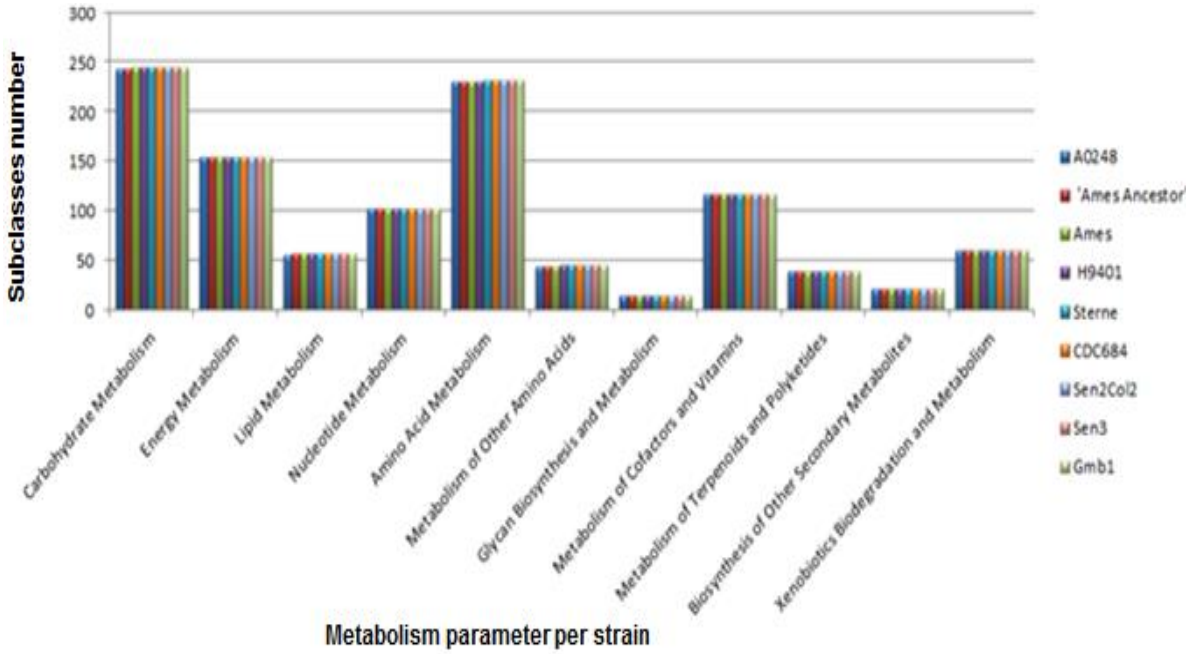

Figure 5. Distribution of metabolisms *B. anthracis* subclasses for various strains.

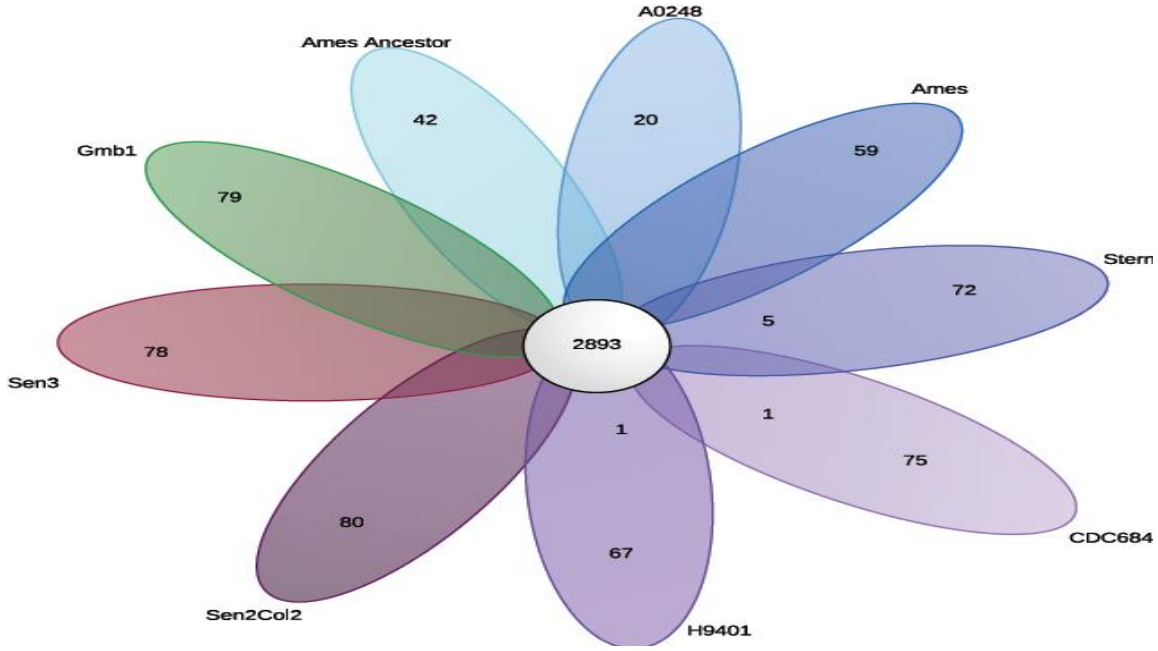

Figure 6. *B. anthracis* pan-genome: Flower pot showing *B. anthracis.* Pan - genome: The number in the center circle is the basic genome. The numbers in the upper part of the petals is the number of accessory genes present in each strain (about 85 in total). The numbers in the lower part of the petals is the number of unique genes in each strain.

niche. It was found out that the three African strains examined belong to lineage A (worldwide lineage), specifically lineage A4, similar to CDC684 and another previously characterized African strain. Pan-genome analysis allowed us to assess the lifestyle of this pathogen and confirmed its allopatric, highly specialized

Figure 7. Hierarchic clustering.

Table 4. Pan-genome for human pathogen strains with colum percentages corresponding to the ratio core/pan genome.

Species	Genome used	Life style	Intracellular	Niche	Pan genome size	Core genome size	%
B .anthracis	9	Allopatric	Non	Animal	47041887	46513801	99
Rickettsia rickettsi	8	Allopatric	Non	Ticks	10129221	100112432	99
Chlamydia trachomatis	20	Allopatric	Non	Human	20960000	20689197	99
Rickettsia prowazeki	8	Allopatric	Non	Human	8888959	8869530	100

lifestyle.

Conflict of Interests

The authors have not declared any conflict of interests.

REFERENCES

Chen F, Mackey AJ, Stoeckert CJ Jr, Roos DS (2006). OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Res. 34(suppl 1):D363-D368.

Chun JH, Hong KJ, Cha SH, Cho MH, Lee KJ, Jeong DH, Yoo CK, Rhie GE (2012). Complete genome sequence of *Bacillus anthracis* H9401, an isolate from a Korean patient with anthrax. J. Bacteriol. 194:4116-4117.

Darling AE, Mau B, Perna NT (2010.) progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS.One. 5:e11147.

Grissa I, Vergnaud G, Pourcel C (2007.) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 35:W52-W57.

Guillemin J (2002). The 1979 anthrax epidemic in the USSR: applied science and political controversy. Proc.Am.Philos.Soc. 146:18-36.

Holden MT, Hsu LY, Kurt K, Weinert LA, Mather AE, Harris SR, Strommenger B, Layer F, Witte W, de Lencastre H, Skov R (2013). A genomic portrait of the emergence, evolution and global spread of a methicillin resistant Staphylococcus aureus pandemic. Genome Res. 23(4):653-664.

Imperiale MJ, Casadevall A (2011). Bioterrorism: lessons learned since the anthrax mailings. MBio. 2:e00232-11.

Kim K, Cheon E, Wheeler KE, Youn Y, Leighton TJ, ParkC, Kim W, Chung SI (2005) Determination of the most closely related bacillus isolates to *Bacillus anthracis* by multilocus sequence typing. Yale J. Biol. Med. 78:1-14.

Kuroda M, Serizawa M, Okutani A, Sekizuka T, Banno S, Inoue S (2010). Genome-wide single nucleotide polymorphism typing method for identification of *Bacillus anthracis* species and strains among B. cereus group species. J. Clin. Microbiol. 48:2821-2829.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007). KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 35:W182-W185.

Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 27:29-34.

Okinaka RT, Price EP, Wolken SR, Gruendike JM, Chung WK, Pearson T, Xie G, Munk C, Hill KK, Challacombe J, Ivins BE (2011). An attenuated strain of *Bacillus anthracis* (CDC 684) has a large chromosomal inversion and altered growth kinetics. BMC Genomics 12:477.

Rasko DA, Worsham PL, Abshire TG, Stanley ST, Bannan JD, Wilson MR, Langham RJ, Decker RS, Jiang L, Read TD, Phillippy AM (2011). *Bacillus anthracis* comparative genome analysis in support of the Amerithrax investigation. Proc. Natl. Acad. Sci. U.S.A 108:5027-5032.

Ravel J, Jiang L, Stanley ST, Wilson MR, Decker RS, Read TD, Worsham P, Keim PS, Salzberg SL, Fraser-Liggett CM, Rasko DA (2009). The complete genome sequence of *Bacillus anthracis* Ames "Ancestor". J. Bacteriol. 191:445-446.

Read TD, Salzberg SL, Pop M, Shumway M, Umayam L, Jiang L, Holtzapple E, Busch JD, Smith KL, Schupp JM, Solomon D (2002). Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*. Science 296:2028-2033.

Read TD, Peterson SN, Tourasse N, Baillie LW, Paulsen IT, Nelson KE, Tettelin H, Fouts DE, Eisen JA, Gill SR, Holtzapple EK (2003). The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. Nature 423:81-86.

Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, Li J, Thiagarajan M, White JA, Quackenbush J (2006). TM4 microarray software suite. Methods Enzymol. 411:134-193.

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A (2003.) TM4: a free, open-source system for microarray data management and analysis. Biotechniques 34:374-378.

Scarlata F, Colletti P, Bonura SI, Trizzino M, Giordano S, Titone L (2010). [The return of anthrax. From bioterrorism to the zoonotic cluster of Sciacca district]. Infez. Med. 18:86-90.

Sheppard SK, Didelot X, Jolley KA, Darling AE, Pascoe B, Meric G, Kelly DJ, Cody A, Colles FM, Strachan NJ, Ogden ID (2013). Progressive genome-wide introgression in agricultural Campylobacter coli. Mol. Ecol. 22:1051-1064.

Song J, Xu Y, White S, Miller KW, Wolinsky M (2005). SNPsFinder--a web-based application for genome-wide discovery of single nucleotide polymorphisms in microbial genomes. Bioinformatics 21:2083-2084.

Sun S, Chen J, Li W, Altinatas I, Lin A, Peltier S, Stocks K, Allen EE, Ellisman M, Grethe J, Wooley J (2011). Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis: the CAMERA resource. Nucleic Acids Res. 39:D546-D55.

Sweeney DA, Hicks CW, Cui X, Li Y, Eichacker PQ (2011). Anthrax infection. Am. J. Respir. Crit. Care Med. 184:1333-1341.

Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 29:22-28.

Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, DeBoy RT(2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc. Natl. Acad. Sci. U.S.A 102:13950-13955.

Wang DB, Tian B, Zhang ZP, Deng JY, Cui ZQ, Yang RF, Wang XY, Wei HP, Zhang XE (2012). Rapid detection of *Bacillus anthracis* spores using a super-paramagnetic lateral-flow immunological detection system. Biosens. Bioelectron. 42:661-667.