



Bacillus

Investigating the genome diversity of *B. cereus* and evolutionary aspects of *B. anthracis* emergence

Leka Papazisi^a, David A. Rasko^{b,1}, Shashikala Ratnayake^a, Geoff R. Bock^a, Brian G. Remortel^a, Lakshmi Appalla^a, Jia Liu^a, Tatiana Dracheva^a, John C. Braisted^a, Shamira Shallom^a, Behnam Jarrahi^a, Erik Snesrud^a, Susie Ahn^a, Qiang Sun^a, Jennifer Rilstone^a, Ole Andreas Økstad^c, Anne-Brit Kolstø^c, Robert D. Fleischmann^a, Scott N. Peterson^{a,*}

ABSTRACT

Here we report the use of a multi-genome DNA microarray to investigate the genome diversity of *Bacillus cereus* group members and elucidate the events associated with the emergence of *Bacillus anthracis* the causative agent of anthrax—a lethal zoonotic disease. We initially performed directed genome sequencing of seven diverse *B. cereus* strains to identify novel sequences encoded in those genomes. The novel genes identified, combined with those publicly available, allowed the design of a “species” DNA microarray. Comparative genomic hybridization analyses of 41 strains indicate that substantial heterogeneity exists with respect to the genes comprising functional role categories. While the acquisition of the plasmid-encoded pathogenicity island (pXO1) and capsule genes (pXO2) represents a crucial landmark dictating the emergence of *B. anthracis*, the evolution of this species and its close relatives was associated with an overall shift in the fraction of genes devoted to energy metabolism, cellular processes, transport, as well as virulence.

© 2011 Elsevier Inc. All rights reserved.



Agenda

- Introduction
- Methods and results
- Conclusion
- Critique



Introduction

Bacillus is a genus of G+ and rod-shaped bacteria.

Three well-known species:

- *B. subtilis* Laboratory model organism
- *B. cereus* (Bc) Causative agent of foodborne illness
- *B. anthracis* (Ba) Causative agent of anthrax

Bc and Ba share 99% 16s rRNA sequence – phylogenetically one group called *Bacillus cereus sensu lato*. Medical reasons for considering them as separate species.



Strains used in the analysis

Table 1
 Strains used in the analysis.

Species	Strain	Alternative designations	Year of isolation	Geographic origin	Source	Type of infection (human isolates)	MLST sequence type (ST) ^a	Plasmid Profile	Main phylogenetic cluster ^b
<i>B. cereus</i>	AH259 ^c	6A1 in BGSC					Not typed		II
<i>B. cereus</i>	AH607 ^c			Norway	Dairy		17		I
<i>B. cereus</i>	AH535 ^c		1994	Norway	Soil (strawberry field)		77		III
<i>B. cereus</i>	AH819 ^c		1995	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH812 ^c		1995	Norway	Human	Periodontitis	38		I
<i>B. cereus</i>	AH1123 ^c			France	Human	Blood	45		I
<i>B. weihenstephanensis</i>	AH1143 ^c			Germany	Dairy (milk)		68		III
<i>B. cereus</i>	ATCC10987		1930	Canada	spores cereal		2		I
<i>B. cereus</i>	ATCC14579 ^c		1916	USA	Air, farmhouse		33		II
<i>B. cereus</i>	AH404			Finland	Dairy		Not typed		III
<i>B. cereus</i>	AH405			Norway	Dairy		Not typed		II
<i>B. cereus</i>	AH407			Finland	Dairy		9		III
<i>B. cereus</i>	AH533		1993	Norway	Soil (strawberry field)		Not typed		IV ^d
<i>B. cereus</i>	AH540		1993	Norway	Soil		Not typed		I
<i>B. cereus</i>	AH557		1993	Norway	Soil		Not typed		I
<i>B. cereus</i>	AH564		1993	Norway	Soil		Not typed		I
<i>B. cereus</i>	AH568		1993	Norway	Soil		Not typed		I
<i>B. cereus</i>	AH597			Norway	Dairy		Not typed		III
<i>B. cereus</i>	AH598			Norway	Dairy		Not typed		I
<i>B. cereus</i>	AH599			Norway	Dairy		Not typed		I
<i>B. cereus</i>	AH601			Norway	Dairy		Not typed		II
<i>B. cereus</i>	AH602			Norway	Dairy		Not typed		III
<i>B. cereus</i>	AH603			Norway	Dairy		Not typed		III
<i>B. cereus</i>	AH604			Norway	Dairy		Not typed		I
<i>B. cereus</i>	AH608			Norway	Dairy		Not typed		I
<i>B. cereus</i>	AH648		1994	Norway	Soil		Not typed		III
<i>B. cereus</i>	AH810		1994	Norway	Human	Periodontitis	36		I
<i>B. cereus</i>	AH811		1995	Brazil	Human	Periodontitis	37		II
<i>B. cereus</i>	AH813		1995	Brazil	Human	Periodontitis	39		I
<i>B. cereus</i>	AH814		1995	Norway	Human	Periodontitis	84		II
<i>B. cereus</i>	AH815		1995	Norway	Human	Periodontitis	85		II
<i>B. cereus</i>	AH816		1995	Norway	Human	Periodontitis	39	pPER272	I
<i>B. cereus</i>	AH817		1995	Norway	Human	Periodontitis	3	pPER272	I
<i>B. cereus</i>	AH818		1995	Brazil	Human	Periodontitis	39		I
<i>B. cereus</i>	AH820		1995	Norway	Human	Periodontitis	39		I
<i>B. cereus</i>	AH823		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH825		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH826		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH827		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH828		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH830		1995	Norway	Human	Periodontitis	41		I
<i>B. cereus</i>	AH834		1995	Norway	Human	Periodontitis	42		I
<i>B. anthracis</i>	Ames		1981	USA	Cow		1	pXO+/pXO2+	I
<i>B. anthracis</i>	Australia94	A0039	1994	Australia	Cattle		1	pXO+/pXO2+	I
<i>B. anthracis</i>	Tsankovskii-1			Former USSR	Livestock vaccine strain		1	pXO+/pXO2+	I
<i>B. anthracis</i>	Voluum	A4088	1935	UK	Cow		1	pXO+/pXO2+	I
<i>B. anthracis</i>	Sterne	34F2	1937	South Africa	Cow (animal vaccine strain)		1	pXO+/pXO2-	I
<i>B. anthracis</i>	Pasteur						1	pXO-/pXO2+	I



Methods – Gene Discovery (GD)

- Identifying strain-specific sequences.
- 7 libraries from uncharacterized Bc strains, randomly cut and attached to glass slides.
- Flip-dye hybridization with reference strain of Ba Ames and query strain.
- Plasmids hybridizing to the query strain, but not Ba Ames were considered strain-specific sequences.



Novel sequences by GD

Table 2
Comparative sequence analysis summary of *B. cereus* novel sequences found from Gene Discovery.

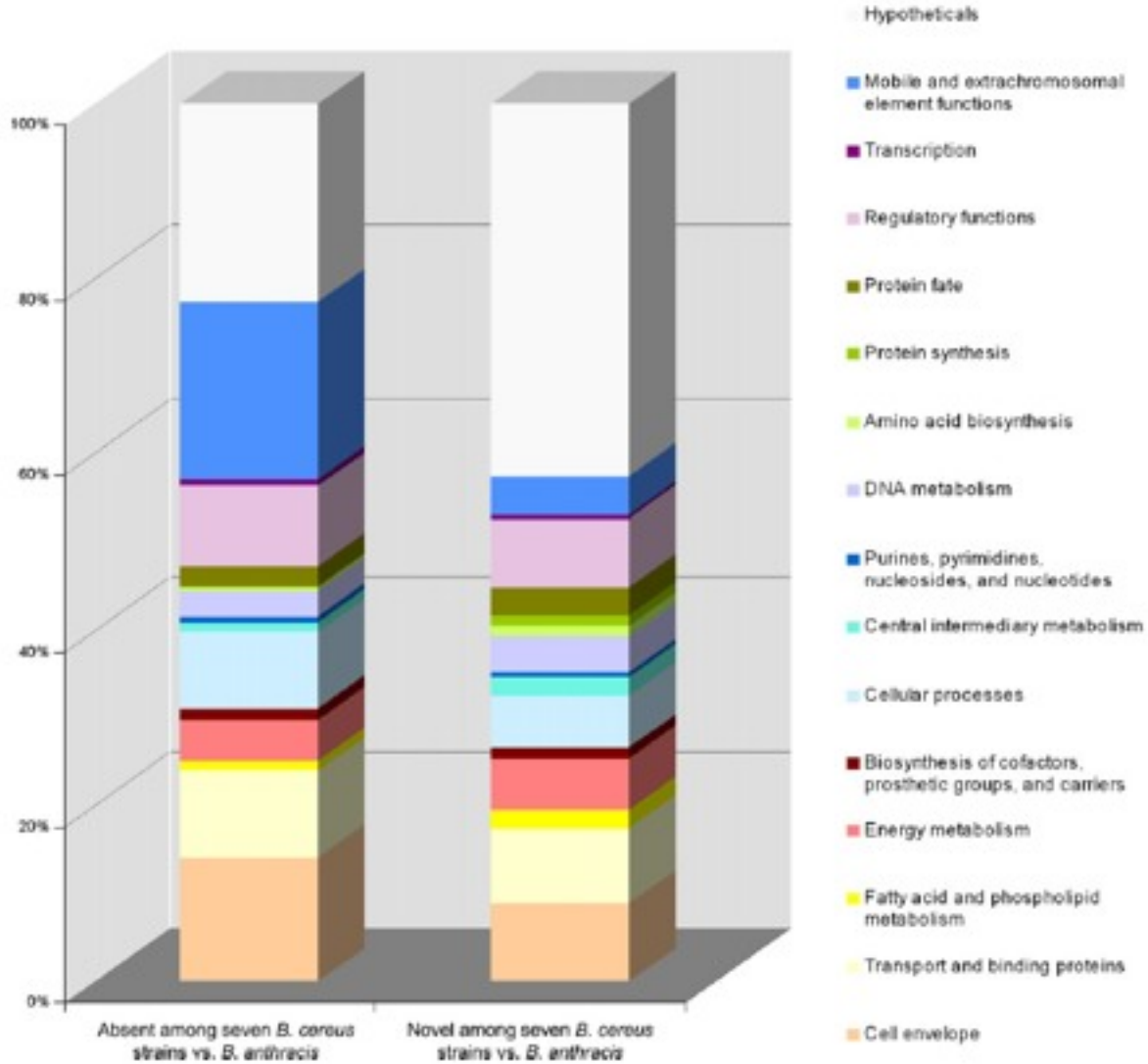
<i>B. cereus</i> strains	AH819	AH607	AH535	AH1123	AH812	AH1143	AH259							
Total bases sequenced/strain	1,076,892 bp	1,237,775 bp	1,500,190 bp	443,218 bp	104,276 bp	1,422,879 bp	177,983 bp							
Fraction relative to the size of an average <i>B. cereus</i> genome	20%	23%	28%	8%	2%	27%	3%							
Total number of annotated features	2041	2204	3138	871	200	2874	383							
<i>Species matches</i>														
<i>B. anthracis</i>	43	2%	122	6%	132	4%	52	6%	7	4%	92	3%	11	3%
<i>B. anthracis</i> pXO1 and pXO2	4	0%	0	0%	0	0%	0	0%	1	1%	2	0%	0	0%
<i>B. cereus</i> 10987	503	25%	510	23%	431	14%	85	10%	34	17%	373	13%	4	1%
pBc10987	121	6%	3	0%	1	0%	1	0%	17	9%	0	0%	0	0%
<i>B. cereus</i> ATCC14579	143	7%	183	8%	427	14%	64	7%	15	8%	394	14%	271	71%
<i>B. cereus</i> (other strain genomes)	200	10%	275	12%	284	9%	192	22%	21	11%	227	8%	12	3%
<i>B. cereus</i> (other strain plasmids)	49	2%	70	3%	112	4%	18	2%	12	6%	73	3%	0	0%
<i>B. thuringiensis</i>	97	5%	245	11%	490	16%	52	6%	7	4%	474	16%	40	10%
<i>B. thuringiensis</i> plasmids	0	0%	6	0%	3	0%	0	0%	0	0%	0	0%	0	0%
<i>Bacillus</i> spp.	1	0.05%	1	0.05%	5	0.16%	0	0%	0	0%	6	0.2%	0	0%
<i>Bacillus</i> spp. plasmids	0	0%	2	0%	0	0%	0	0%	0	0%	1	0%	0	0%
<i>Bacillus</i> spp. phages	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
Overall number of features similar to <i>Bacillus</i> species by BlastN	1161	57%	1417	64%	1885	60%	464	53%	114	57%	1642	57%	338	88%
Remaining relatively novel features (by BlastN)	880	43%	787	36%	1253	40%	407	47%	86	43%	1232	43%	45	12%
Truly unique features from the relatively novel gene set ^b	452	22%	312	14%	887	28%	324	37%	71	36%	1054	37%	19	5%

^a Fraction from the total number of annotated features in the corresponding strain.

^b Remaining pORFs whose amino acid sequence shows no significant homology to proteins in databases by BlastP.

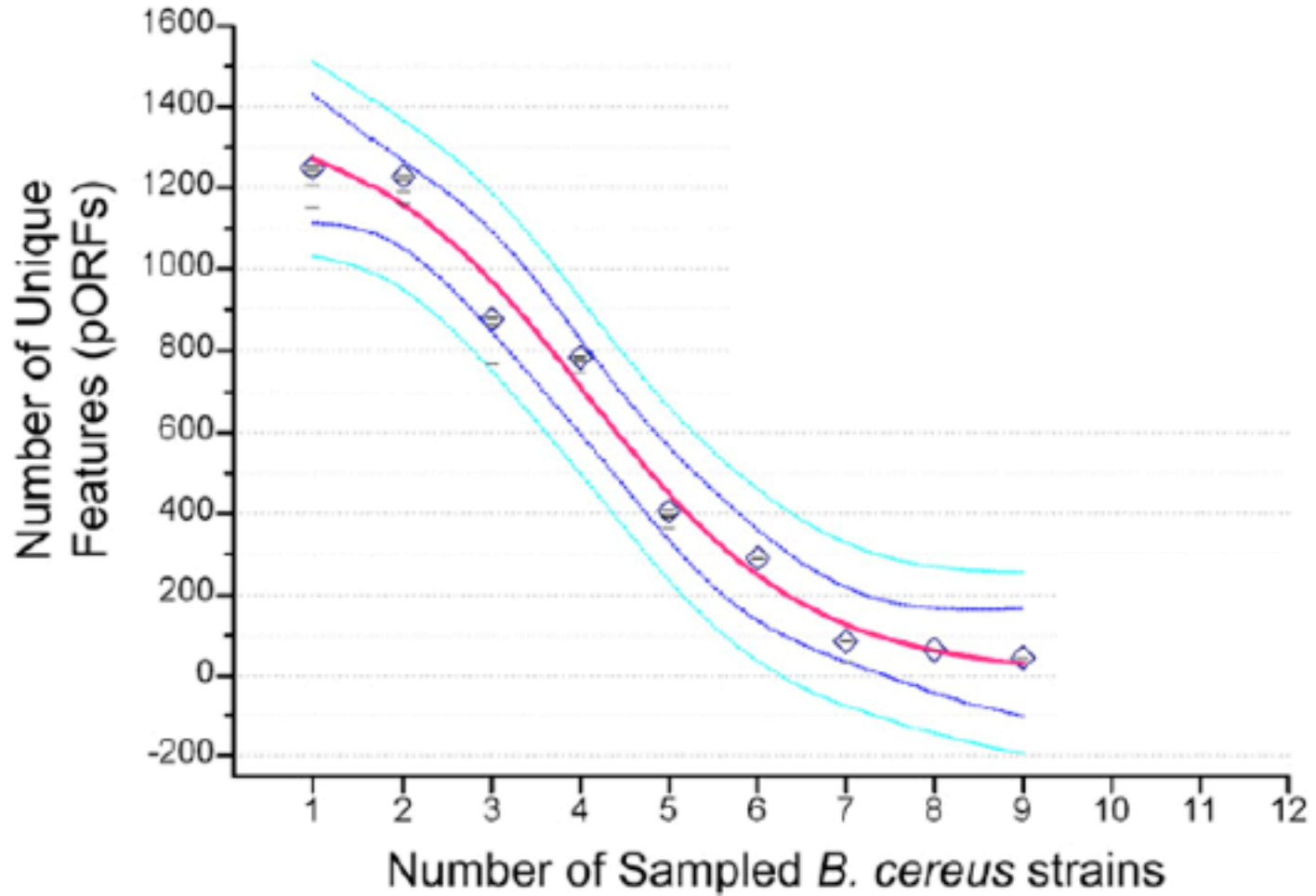


Comparative summary in %





Number of novel pORFs



Strains used in the analysis

Table 1
Strains used in the analysis.

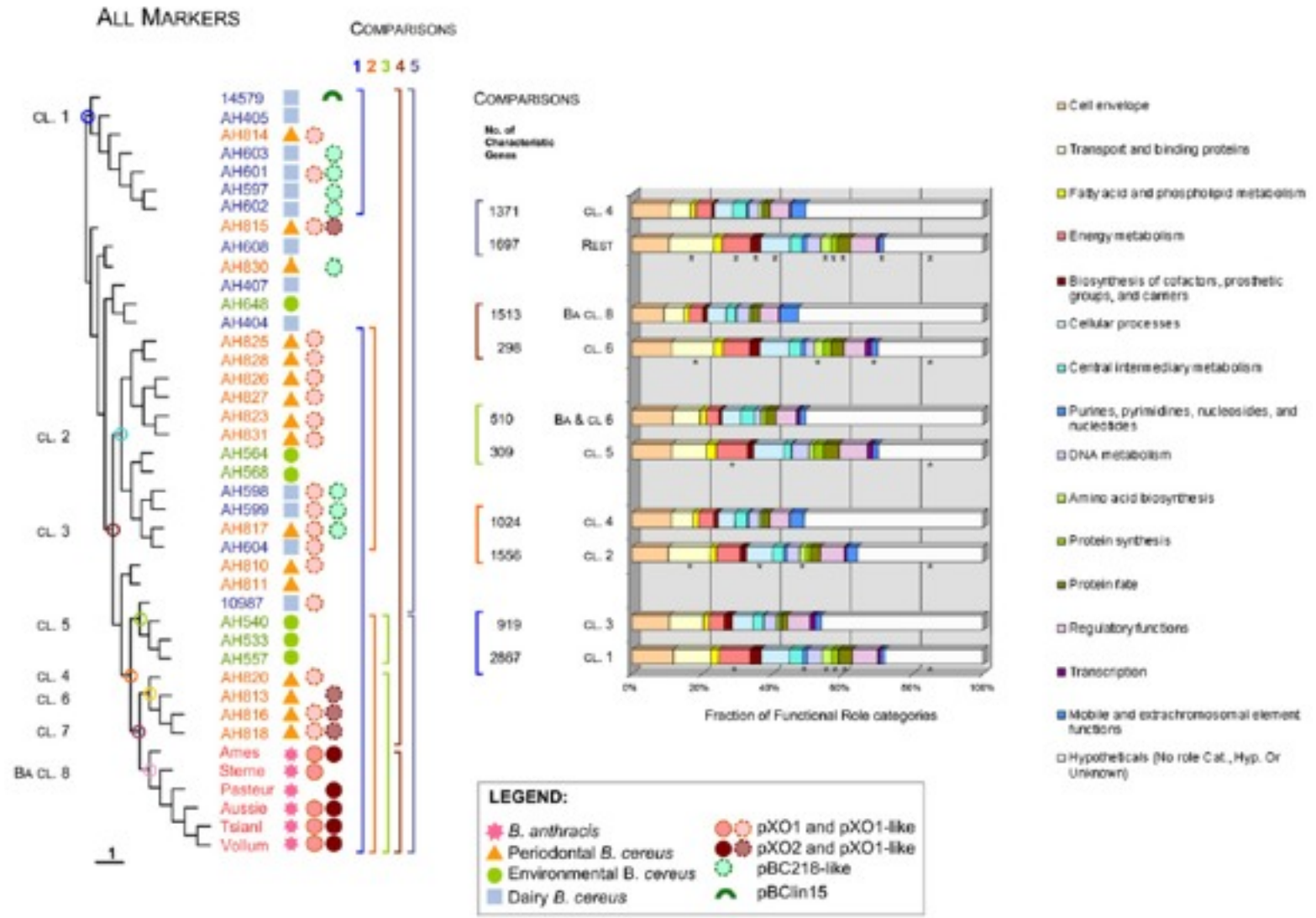
Species	Strain	Alternative designations	Year of isolation	Geographic origin	Source	Type of infection (human isolates)	MLST sequence type (ST) ^a	Plasmid Profile	Main phylogenetic cluster ^b
<i>B. cereus</i>	AH259 ^c	6A1 in BGSC					Not typed		II
<i>B. cereus</i>	AH607 ^c			Norway	Dairy		17		I
<i>B. cereus</i>	AH535 ^c		1994	Norway	Soil (strawberry field)		77		III
<i>B. cereus</i>	AH819 ^c		1995	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH812 ^c		1995	Norway	Human	Periodontitis	38		I
<i>B. cereus</i>	AH1123 ^c			France	Human	Blood	45		I
<i>B. weihenstephanensis</i>	AH1143 ^c			Germany	Dairy (milk)		68		III
<i>B. cereus</i>	ATCC10987		1930	Canada	spores cereal		2		I
<i>B. cereus</i>	ATCC14579 ^c		1916	USA	Air, farmhouse		33		II
<i>B. cereus</i>	AH404			Finland	Dairy		Not typed		III
<i>B. cereus</i>	AH405			Norway	Dairy		Not typed		II
<i>B. cereus</i>	AH407			Finland	Dairy		9		III
<i>B. cereus</i>	AH533		1993	Norway	Soil (strawberry field)		Not typed		IV ^d
<i>B. cereus</i>	AH540		1993	Norway	Soil		Not typed		I
<i>B. cereus</i>	AH557		1993	Norway	Soil		Not typed		I
<i>B. cereus</i>	AH564		1993	Norway	Soil		Not typed		I
<i>B. cereus</i>	AH568		1993	Norway	Soil		Not typed		I
<i>B. cereus</i>	AH597			Norway	Dairy		Not typed		III
<i>B. cereus</i>	AH598			Norway	Dairy		Not typed		I
<i>B. cereus</i>	AH599			Norway	Dairy		Not typed		I
<i>B. cereus</i>	AH601			Norway	Dairy		Not typed		II
<i>B. cereus</i>	AH602			Norway	Dairy		Not typed		III
<i>B. cereus</i>	AH603			Norway	Dairy		Not typed		III
<i>B. cereus</i>	AH604			Norway	Dairy		Not typed		I
<i>B. cereus</i>	AH608			Norway	Dairy		Not typed		I
<i>B. cereus</i>	AH648		1994	Norway	Soil		Not typed		III
<i>B. cereus</i>	AH810		1994	Norway	Human	Periodontitis	36		I
<i>B. cereus</i>	AH811		1995	Brazil	Human	Periodontitis	37		II
<i>B. cereus</i>	AH813		1995	Brazil	Human	Periodontitis	39		I
<i>B. cereus</i>	AH814		1995	Norway	Human	Periodontitis	84		II
<i>B. cereus</i>	AH815		1995	Norway	Human	Periodontitis	85		II
<i>B. cereus</i>	AH816		1995	Norway	Human	Periodontitis	39	pPER272	I
<i>B. cereus</i>	AH817		1995	Norway	Human	Periodontitis	3	pPER272	I
<i>B. cereus</i>	AH818		1995	Brazil	Human	Periodontitis	39		I
<i>B. cereus</i>	AH820		1995	Norway	Human	Periodontitis	39		I
<i>B. cereus</i>	AH823		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH825		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH826		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH827		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH828		1996	Norway	Human	Periodontitis	40		I
<i>B. cereus</i>	AH830		1995	Norway	Human	Periodontitis	41		I
<i>B. cereus</i>	AH831		1996	Norway	Human	Periodontitis	42		I
<i>B. anthracis</i>	Ames		1981	USA	Cow		1	pXO+/pXO2+	I
<i>B. anthracis</i>	Australia94	A0039	1994	Australia	Cattle		1	pXO+/pXO2+	I
<i>B. anthracis</i>	Tsankovskii-1			Former USSR	Livestock vaccine strain		1	pXO+/pXO2+	I
<i>B. anthracis</i>	Voluum	A4088	1935	UK	Cow		1	pXO+/pXO2+	I
<i>B. anthracis</i>	Sterne	34F2	1937	South Africa	Cow (animal vaccine strain)		1	pXO+/pXO2-	I
<i>B. anthracis</i>	Pasteur						1	pXO-/pXO2+	I



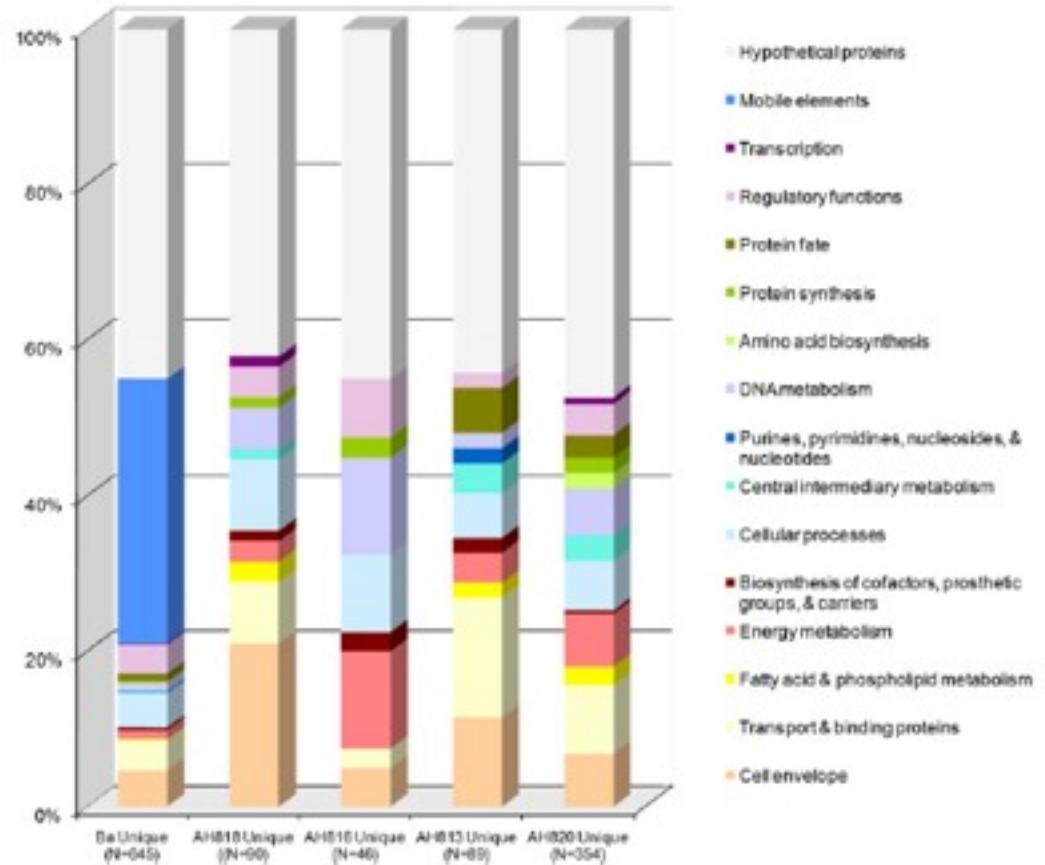
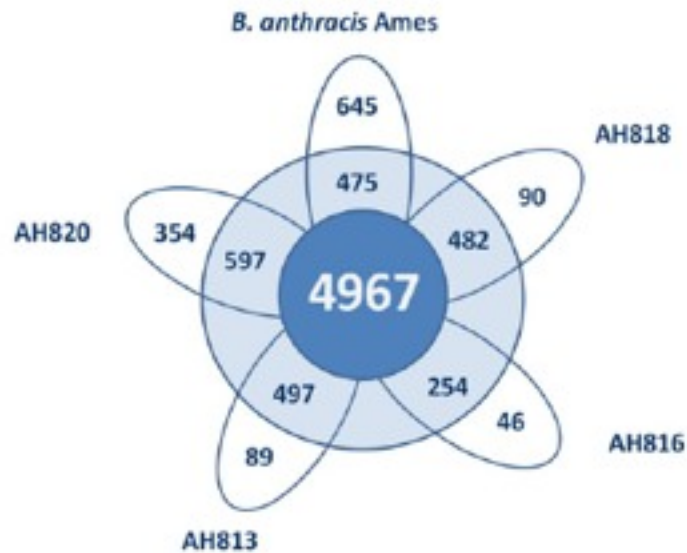
Methods – Comparative Genomic Hybridization (CGH)

- 70-mer oligonucleotide DNA microarray
- 29,682 ORFs
- 35 Bc and six Ba strains were tested

Gene clustering and comparisons



Genes of clade 7



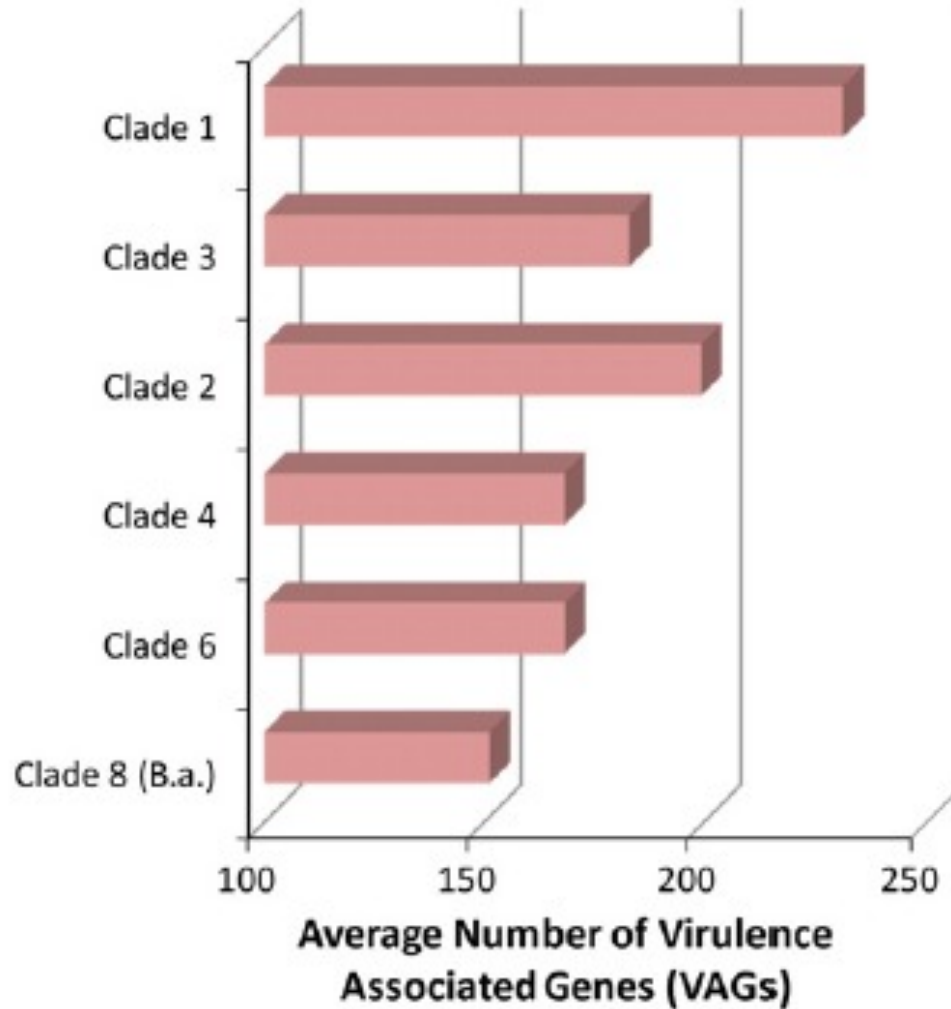


Characteristic VAGs of *B. anthracis*

Locus	Common name	Characteristic for	
		Clade 4 genomes vs. other <i>B. cereus</i>	<i>B. anthracis</i> vs. <i>B. cereus</i>
BA0552	Internalin, putative	(+) ^a	+
BA1489	Superoxide dismutase	+	+
BA1760	Cobalamin synthesis protein, putative	+	+
BA1902	Multicopper oxidase family protein	+	+
BA1925	Cobalamin synthesis protein/P47K family protein	+	+
BA1981 ^b	Siderophore biosynthesis protein, putative	+	+
BA1982 ^b	Siderophore biosynthesis protein, putative	+	+
BA1983 ^b	AMP-binding protein	+	+
BA1984 ^b	Hypothetical protein	+	+
BA1985 ^b	Hypothetical protein	+	+
BA1986 ^b	Conserved hypothetical protein	+	+
BA1992	Phospholipase, putative	+	+
BA2222	Alcohol dehydrogenase, iron-containing	+	+
BA2372 ^b	Non-ribosomal peptide synthetase Dhhf	+	+
BA2588	Alcohol dehydrogenase, zinc-containing	+	+
BA2642	Cobalt transport protein	+	+
BA2730	Neutral protease	+	+
BA3100	Copper homeostasis protein CutC, putative	+	+
BA3131	Alcohol dehydrogenase, zinc-containing	+	+
BA3189	Manganese ABC transporter, substrate-binding protein/adhesin	(+) ^a	+
BA3269	Iron-sulfur cluster-binding protein	+	+
BA3299	Microbial collagenase, putative	+	+
BA3307 ^b	L-Serine dehydratase, iron-sulfur-dependent, alpha subunit	+	+
BA3308 ^b	L-Serine dehydratase, iron-sulfur-dependent, beta subunit	+	+
BA3435 ^b	Alcohol dehydrogenase, zinc-containing	+	+
BA3442	Neutral protease	+	+
BA3515	Alcohol dehydrogenase, zinc-containing, authentic point mutation	+	+
BA3584	Microbial collagenase, putative	+	+
BA3703	Phospholipase/carboxylesterase family protein	+	+
BA3922	Zinc protease, insulinase family	+	+
BA4766 ^b	Iron compound ABC transporter, iron compound-binding protein	+	+
BA4767 ^b	Iron compound ABC transporter, permease protein	+	+
BA4784 ^b	Iron compound ABC transporter, ATP-binding protein	+	+
BA5701	Channel protein, hemolysin III family	+	+
BXA0142	Calmodulin-sensitive adenylate cyclase, Cya		+ ^c
BXA0146	Transcriptional activator AtoA, (pX01-119)		+ ^c
BXA0163	Protective antigen-related protein, (pX01-111)		+ ^c
BXA0164	Protective antigen, PagA		+ ^c
BXA0172	Lethal factor, Lef		+ ^c
BXA0197	Zinc-binding lipoprotein AdcA domain protein, (pX01-130)		+ ^c
BX00060	Capsule synthesis trans-acting positive regulator, putative, (pX02-53)		+
BX00062 ^b	Hypothetical protein, CapI, (pX02-54)		+ ^c
BX00063 ^b	Gamma-glutamyltranspeptidase, capD, (pX02-55)		+ ^c
BX00064 ^b	Capsule biosynthesis protein CapA, (pX02-56)		+
BX00065 ^b	Capsule biosynthesis protein CapC, (pX02-57)		+ ^c
BX00066 ^b	Capsule biosynthesis protein CapB, (pX02-58)		+ ^c
BX00084	Capsule synthesis trans-acting positive regulator, CapR, (pX02-64)		+ ^c



VAGs of the different clades





Conclusions

- *B. cereus sensu lato* share 99 % identity in 16s rDNA, however GD demonstrated up to 37 % novel genes in one *B. cereus* strain out of 7 tested.
- *B. anthracis* differ from the *B. cereus* mainly in the higher genomic percentage of mobile elements.
- *B. cereus sensu lato* can be clustered into 8 clades. Interestingly clade 8 (Ba) and 6 show high similarity.
- The pathogenic *B. anthracis* is associated with a fixation of a specific set of VAGs.
- The evolution of *B. anthracis* has driven towards deeper dependency on the host for energy and



Critique Points

- Methods – whole genome sequencing instead of microarray?
- Supplementary figures and tables
- Relative results – other ways to illustrate results



Critique Points

- Elucidates the gene composition of the different isolates
- Leaves room for further investigation of the virulens factors of *B. anthracis*



Thanks