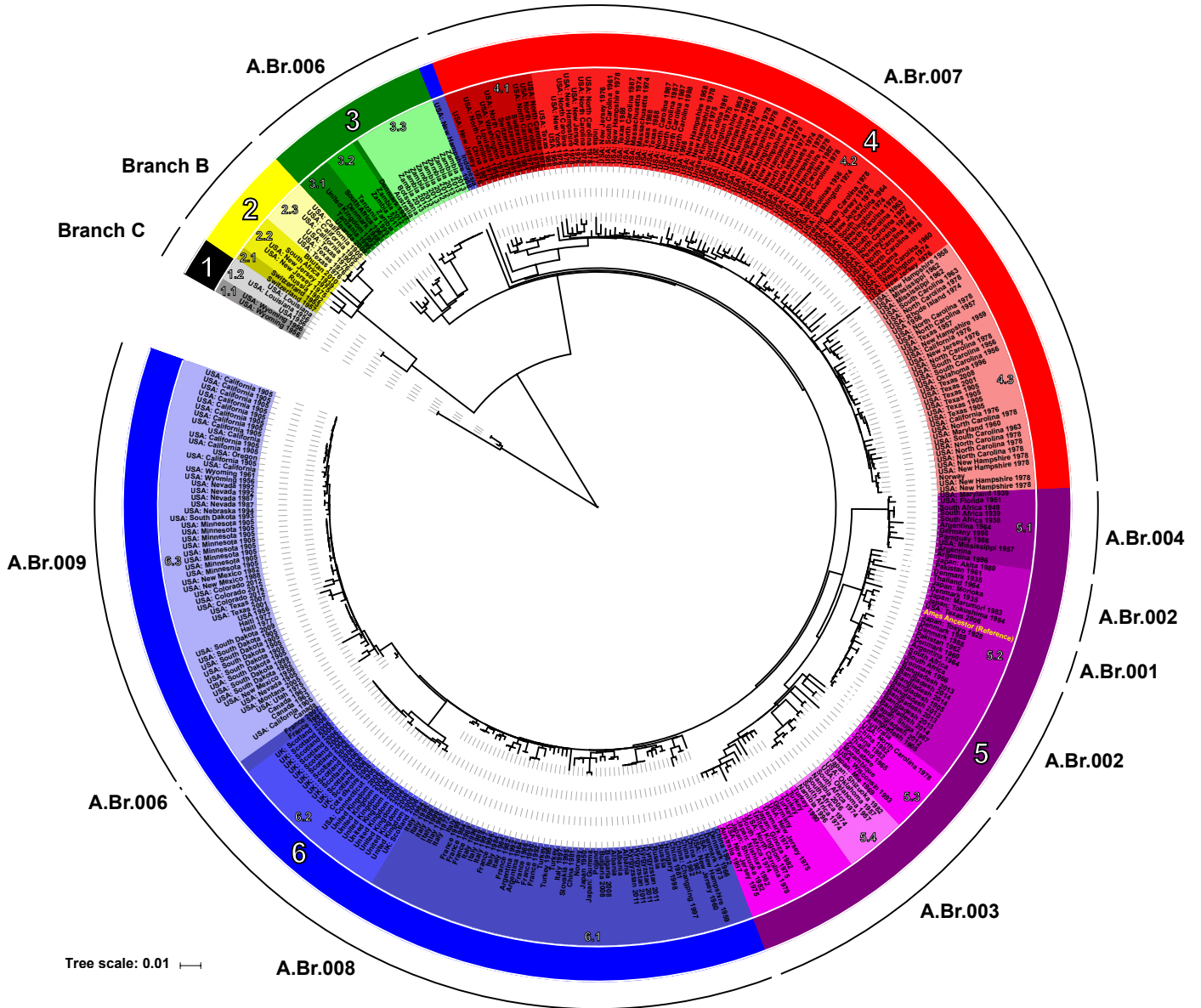
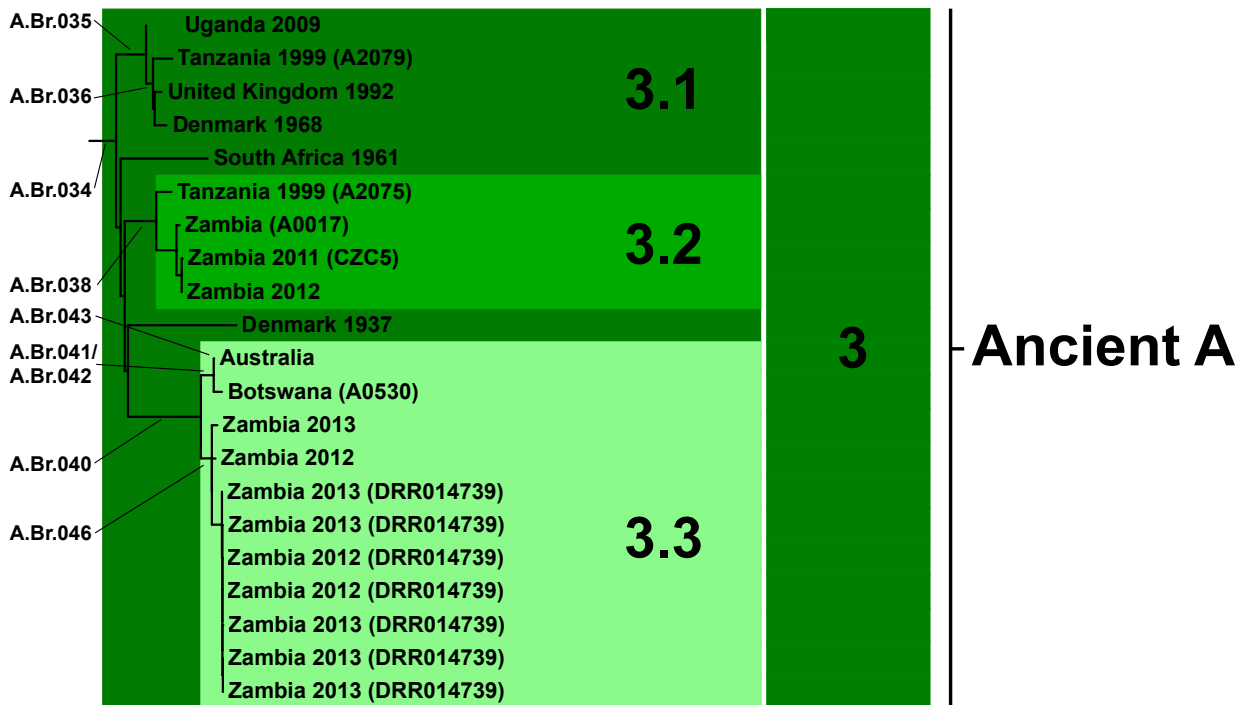
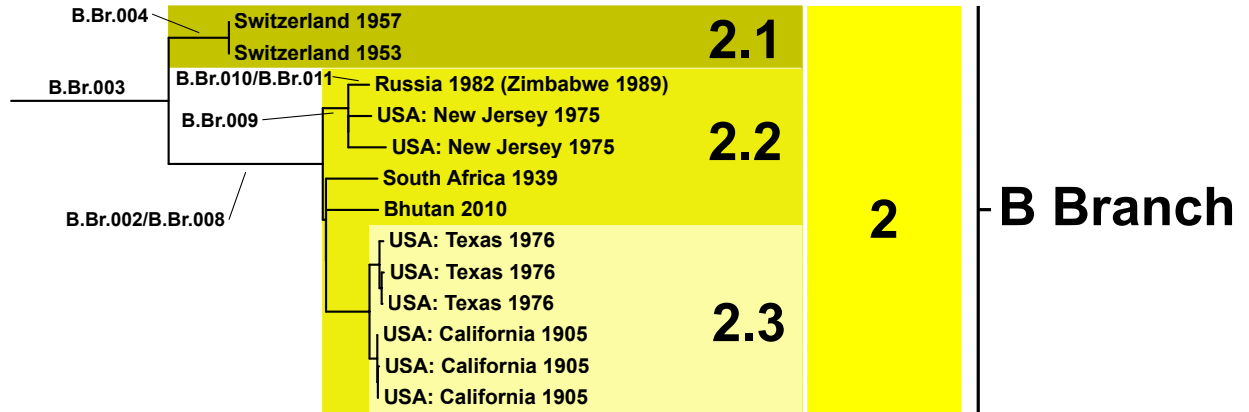
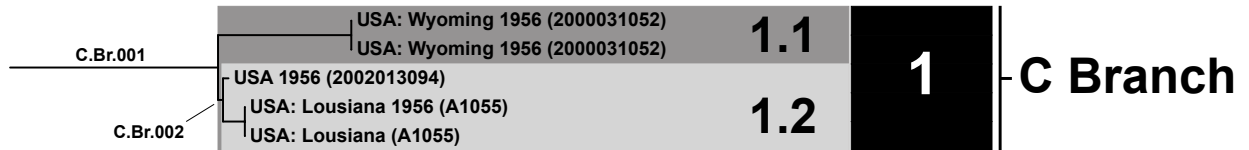


Figure S1. Whole-chromosome tree of 356 global *B. anthracis* isolates from Bruce et al., 2020. Primary clusters 1–6 are designated in outer colored ring. Primary clusters are further divided into nested clades indicated with shaded tip labels. The outer black lines with labels represent the major lineages defined in Van Ert et al., 2007. The location of the *B. anthracis* reference genome (NCBI: accession AE017334.2) within cluster 5 is labeled in yellow font. Each primary cluster is then broken out on its own, with branch labels from Sahl et al., 2016.





	USA: New Jersey 1976	
	USA: China 1981 (A0615)	
	USA: North Carolina 1978 (A0615)	
	USA: Louisiana 1956 (A0615)	
	USA: Washington 1974 (A0615)	
	USA: North Carolina 1978 (K1129)	4.1
A.Br.050	Switzerland 1981 (K1129)	
	Switzerland 1981 (K1129)	
	Switzerland 1981 (K1129)	
A.Br.047/ A.Br.007	USA: North Carolina 1978 (K1129)	
	USA: North Carolina 1978 (K1129)	
	USA: North Carolina 1978 (K1129)	
	USA: Texas 1988	
	USA 1956	
	USA: New York 1976	
	USA: North Carolina 1978	
	USA: New Hampshire 1978	
	USA: New Jersey 1976	
	USA: North Carolina 1978	
A.Br.051	USA: North Carolina 1978	
	India 2017	
	USA: New Jersey 1976 (SK102)	
	USA: South Carolina 1961	
	USA: New Hampshire 1978	
	USA: Texas 1988	
	USA: North Carolina 1987	
	USA: Massachusetts 1974	
	USA: Massachusetts 1974	
	USA: Texas 1988	
	USA: Texas 1988	
	USA: North Carolina 1987	
	USA: North Carolina 1987	
	USA: North Carolina 1987	
	USA: North Carolina 1988	
	USA 1956	
	USA: New Hampshire 1958	
	USA: New Hampshire 1978	
	United Kingdom 1975	
	USA: South Carolina 1961	
	USA: Washington 1975	
	USA: New Hampshire 1958	
	USA: New Hampshire 1958	
	USA: New Hampshire 1958	
	USA: Washington 1974	
	USA: New Hampshire 1978 (Smith1013)	
	USA: New Hampshire 1978	
	USA: Washington 1974	
	USA: Washington 1974	
	USA: New Hampshire 1978	4.2
	USA: New Hampshire 1978	
	USA: North Carolina 1978	
	USA: Rhode Island 1974	
	USA: Massachusetts 1974	
	USA: New Hampshire 1978	
	USA: New Hampshire 1978	
	USA: North Carolina 1987	
	USA: North Carolina 1978	
	USA 1956 (2000031709)	
	USA: Carolinas 1956 (2000031709)	
	USA: Washington 1974	
	Ireland (A0390)	
A.Br.052	USA: North Carolina 1978	
	USA: New Jersey 1976	
	USA: New Jersey 1976	
	USA: South Carolina 1964	
	USA: Washington 1974	
A.Br.054	USA: North Carolina 1978	
	USA: South Carolina 1963	
	USA: North Carolina 1957	
	USA: Pennsylvania 1974	
	USA: North Carolina 1961	
A.Br.053	USA: North Carolina 1978	
	USA: Alabama (2000031008)	
	USA: South Carolina 1960	
	USA: Washington 1974	
	USA: New Jersey 1976	
	USA: New Hampshire 1958	
	USA: Mississippi 1962	
	USA: Mississippi 1962	
	USA: South Carolina 1963	
	USA: North Carolina 1978	
	USA: Rhode Island 1974	
	USA 1956	
	USA: North Carolina 1978	
	USA: North Carolina 1957	
A.Br.055	USA: Texas 1957	
A.Br.056	USA: New Hampshire 1959	
A.Br.058	USA: California 1976	
	USA: New Jersey 1976	
	USA: North Carolina 1978	
	USA: South Carolina 1956 (A0860)	
	USA: South Carolina 1956	
	USA: Oklahoma 1996	4.3
	USA: Texas 2008	
	USA: Texas 2001	
	USA: Texas 1905	
	USA: Texas 1905	
	USA: Texas 1905	
	USA: Texas 1905	
	USA: California 1976	
	USA: North Carolina 1978	
A.Br.059	USA: Maryland 1960 (A1063)	
	USA: South Carolina 1963	
	USA: North Carolina 1978	
A.Br.062	USA: North Carolina 1978	
	USA: North Carolina 1978	
	USA: New Hampshire 1978 (A0363)	
	USA: New Hampshire 1978 (A0363)	
	Norway (A0363)	
	USA: New Hampshire 1978 (A0363)	
	USA: New Hampshire 1978 (A0363)	

4 - Vollum

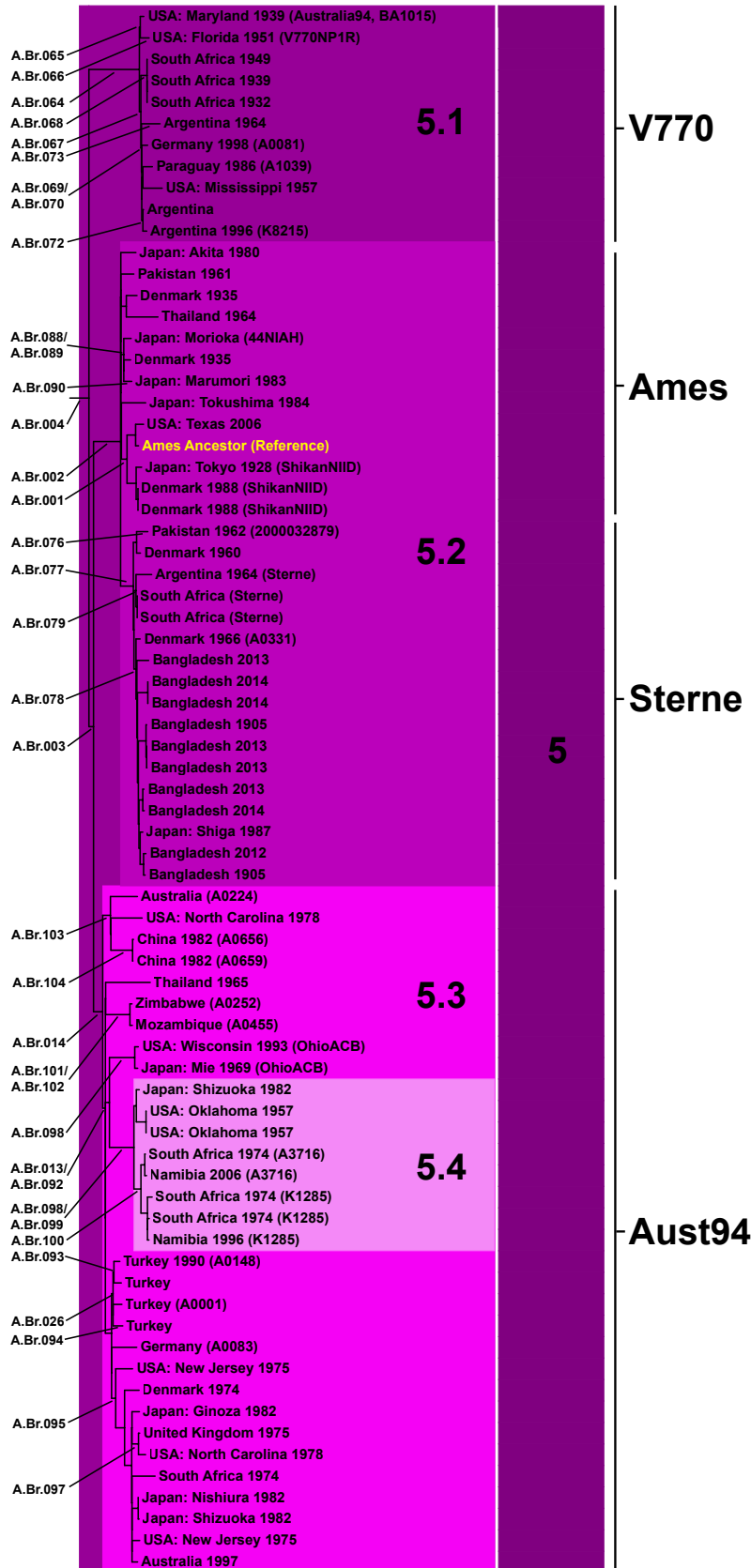


Table S1. PHASTER's phage completeness score calculation (from Zhou et al., 2011). A prophage region is considered to be incomplete if its completeness score is less than 60, questionable if the score is between 60 and 90, and intact if the score is above 90.

	Scenario A	Scenario B	Scenario C
Number of nucleotides (# bases)	–	(# bases in the region/ # bases in the related phage) × 100	+10 if # bases >30 kb
Number of genes (# genes)	–	(# genes in the region/ # genes in the related phage) × 100	+10 if # genes >40
Cornerstone genes ^a	–	–	+10 for each cornerstone gene
Phage-like genes	–	–	+10 if occupies 70% or more of the region
Final scores	150	Sum of above	Sum of above

The completeness score is calculated for three different scenarios: (A) the region contains all genes of a known phage. (B) >50% of the genes in the region are related to a known phage. (C) <50% of the genes in the region are related to a known phage.

^a Cornerstone genes are identified key phage structural genes (using keywords such as ‘capsid’, ‘head’, ‘plate’, ‘tail’, ‘coat’, ‘portal’ and ‘holin’) and phage DNA regulation genes (such as ‘integrase’, ‘transposase’ and ‘terminase’) and phage function genes (such as ‘lysin’ and ‘bacteriocin’).

Table S2. Variables included in the Random Forest classification. See methods for variable sources. Definitions for antimicrobial resistance can be found in Supplemental 1 and virus lineages in Supplemental 2.

Biosample	Antimicrobial resistance gene	Phage sequence
Continent of collection	<i>fosB</i>	Staphylococcus phage vB_SepS_SEP9
Isolation source (host, environment, or industry)*	<i>bla1</i>	Bacillus phage phi4B1
	<i>bla2</i>	Bacillus phage phBC6A52
	<i>vmIR</i>	Bacillus virus Wbeta
	<i>bclI</i>	Bacillus phage Gamma
	<i>tem-116</i>	Enterobacteria phage phiX174
	<i>cfrC</i>	Bacillus phage phiCM3
	<i>dfrG</i>	Bacillus virus AP50
	<i>oxa-59</i>	

*categorical variable based on whether the isolate was collected from a host (e.g. animal), the environment (e.g. soil sample), or industry (e.g. textile factory or animal processing facility)

Table S3. Classification, number of isolates (N), and frequency of each antimicrobial resistance gene across the global whole-genome dataset.

Cluster.clade	Group	N	<i>mphL</i>	<i>fosB</i>	<i>bla1</i>	<i>bla2</i>	<i>vmrR</i>	<i>bcII</i>	<i>tem-116</i>	<i>cfrC</i>	<i>dfxG</i>	<i>oxa-59</i>
1.1	C	2	1.00	0.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
1.2		3	1.00	0.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
2.1	B	2	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
2.2		5	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
2.3		6	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
3.1	Ancient A	6	1.00	1.00	1.00	0.33	1.00	0.00	0.00	0.00	0.00	0.00
3.2		4	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
3.3		11	1.00	1.00	1.00	1.00	1.00	0.00	0.27	0.00	0.09	0.00
4.1	Vollum	12	1.00	1.00	1.00	1.00	0.75	0.00	0.00	0.00	0.00	0.00
4.2		63	1.00	0.98	1.00	0.98	1.00	0.02	0.00	0.00	0.00	0.00
4.3		35	1.00	1.00	0.97	0.97	1.00	0.00	0.00	0.00	0.00	0.00
5.1	V770	11	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
5.2	Ames/Sterne	29	1.00	1.00	0.97	0.93	0.00	0.07	0.00	0.00	0.00	0.00
5.3	Aust 94	24	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.04	0.00	0.00
5.4		8	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
6.1	TEABr.008/011	57	1.00	1.00	1.00	0.50	0.98	0.05	0.00	0.00	0.00	0.02
6.2	TEABr.011	22	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
6.3	WNA	56	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00

Table S4. Classification, number of isolates (*N*), and frequency of each phage sequence across the global whole-genome dataset. Frequency values are listed as intact/questionable/incomplete (i/q/i). Dashes (-) represent absence of the phage sequence.

Cluster. clade	Group	<i>N</i>	<i>Bacillus</i> virus 1	<i>Bacillus</i> phage PFEFR-5	<i>Staphylococcus</i> phage vB_SepS_SEP9	<i>Bacillus</i> phage phBC6A52	<i>Bacillus</i> phage phi4B1	<i>Bacillus</i> phage phi4J1	<i>Enterobacteria</i> phage phiX174	<i>Bacillus</i> phage Gamma	<i>Bacillus</i> virus Wbeta	<i>Bacillus</i> phage phiCM3	<i>Bacillus</i> virus AP50
1.1	C	2	-/1.00/-	-/1.00/-	-/1.00/-	0.50/-/-	-	-	-	-	-	-	-
1.2		3	-/1.00/-	/0.33/0.67	-/1.00/-	0.33/-/-	0.33/-/-	-	-/0.33/-	0.33/-/-	-	-	-
2.1	B	2	-/1.00/-	/0.50/0.50	-/1.00/-	1.00/-/-	-/-/0.33	-	-	-	-	-	-
2.2		5	0.80/0.20/-	-/1.00/-	-/1.00/-	0.60/-/-	-	-/-/0.20	-	-	-	-	-
2.3		6	0.83/0.17/-	-/1.00/-	-/1.00/-	0.50/-/-	-	-/-/0.17	-	-	-	-	-
3.1	Ancient A	6	0.83/0.17/-	-/1.00/-	-/1.00/-	0.17/-/-	-	-/-/0.17	-	-	-	-	-
3.2		4	0.75/0.25/-	-/1.00/-	-/1.00/-	-	-/-/0.25	-	-	-	-	-	-
3.3		11	0.91/0.09/-	-/1.00/-	-/1.00/-	0.18/-/-	-	-/-/0.09	-	-	-	-	-
4.1	Vollum	12	0.92/0.08/-	-	-/0.92/-	0.42/-/-	-/-/0.25	-	-	-	-	-	-
4.2		63	1.00/-/-	-/1.00/-	-/1.00/-	0.29/-/-	-	-	-/0.08/-	-	-	-	-
4.3		35	0.97/0.03/-	/0.97/0.03	-/1.00/-	0.34/-/-	-	-/-/0.03	-	-	0.03/-/-	-	-
5.1	V770	11	1.00/-/-	-/1.00/-	-/1.00/-	0.27/-/-	-	-	-/0.09/-	0.09/-/-	-	-	-
5.2	Ames/Sterne	29	0.90/0.10/-	-	-/1.00/-	0.17/-/0.03	-/-/0.10	-	-	-	-	-	-
5.3	Aust 94	24	1.00/-/-	-/1.00/-	-/1.00/-	0.21/-/-	-/-/0.04	-	-	-	-	-	-
5.4		8	1.00/-/-	-/1.00/-	-/1.00/-	0.13/0.13/-	-	-	-	-	-	-	-
6.1	TEABr.008/011	57	1.00/-/-	/0.98/0.02	-	-/1.00/-	0.25/-/-	-/-/0.02	-	-	0.02/-/0.02	-	-
6.2	TEABr.011	22	0.91/0.09/-	0.05/0.95/	-	-/1.00/-	0.18/-/-	-/-/0.23	-	-	-	-	-
6.3	WNA	56	0.98/0.02/-	-/1.00/-	-/1.00/-	0.20/-/-	-	-/-/0.02	0.04/0.04/	-	0.02/-/-	-	-/-/0.02 0.02/-/-

Table S5. Prophage sequence names, accession numbers, and their full lineages from the National Center for Biotechnology Information (NCBI) database.

Name	Accession	Lineage
Bacillus virus 1	NC_009737	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Myoviridae; Svinavirus
Bacillus phage PFEFR-5	NC_031055	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Siphoviridae; unclassified Siphoviridae
Staphylococcus phage vB_SepS_SEP9	NC_023582	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Siphoviridae; Sextaecvirus; Staphylococcus virus SEP9
Bacillus phage phBC6A52	NC_004821	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Siphoviridae; unclassified Siphoviridae
Bacillus phage phi4B1	NC_028886	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Siphoviridae; unclassified Siphoviridae
Bacillus phage phi4J1	NC_029008	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Siphoviridae; unclassified Siphoviridae
Enterobacteria phage phiX174	NC_001422	Viruses; Monodnaviria; Sangervirae; Phixviricota; Malgrandaviricetes; Petitvirales; Microviridae; Bullavirinae; Sinsheimervirus; Escherichia virus phiX174
Bacillus phage Gamma	NC_007458	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Siphoviridae; Wbetavirus; unclassified Wbetavirus
Bacillus virus Wbeta	NC_007734	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Siphoviridae; Wbetavirus
Bacillus phage phiCM3	NC_023599	Viruses; Duplodnaviria; Heunggongvirae; Uroviricota; Caudoviricetes; Caudovirales; Siphoviridae; unclassified Siphoviridae
Bacillus virus AP50	NC_011523	Viruses; Varidnaviria; Bamfordvirae; Preplasmiviricota; Tectiliviricetes; Kalamavirales; Tectiviridae; Betatectivirus

Table S6. Confusion Matrix from the random forest classifier for population clusters 1 through 6. Variables are listed in Supplementary 4.

		Reference data (the cluster results got from the DNA data; truth)								
Classified data (the results from RF; predicted)	Cluster	1	2	3	4	5	6	Total	User's accuracy (%)	
	1	0	0	0	0	0	0	0	0	-
	2	0	0	0	0	0	0	0	-	
	3	0	0	3	0	0	0	3	100	
	4	1	1	0	19	0	1	22	86.4	
	5	0	0	0	2	14	0	16	87.5	
	6	0	1	1	3	0	26	31	83.9	
	Total	1	2	4	24	14	27			
	Producer's accuracy (%)	0	0	75	79.2	100	96.3			
Kappa = 0.799 Overall accuracy = 86.1%										

Table S7. Potentially high impact SNPs and their predicted effects. Positions are relative to the Ames Ancestor reference genome (NCBI: accession AE017334.2).

Population Cluster	Position	Ref	Alt	Comparative effect	Gene/Product	
1	399920	G	A	stop gained	putative ABC transporter, permease protein	
	923493	C	T	stop gained	oligopeptide ABC transporter, oligopeptide-binding protein	
	1260604	C	T	stop gained	DNA-binding response regulator	
	1292469	C	A	stop gained	stage 0 sporulation regulatory protein	
	1319426	G	A	stop gained	dltD	
	1505215	C	T	stop gained	conserved hypothetical protein	
	1656148	T	A	stop gained	fumarate hydratase, class II	
	3057925	G	A	stop gained	conserved hypothetical protein	
	3378225	G	A	stop gained	hypothetical protein	
	3550604	C	T	stop gained	sensory box/GGDEF family protein	
	3987377	C	T	stop gained	putative ribosomal-protein-alanine acetyltransferase	
	5010067	T	A	stop gained	glycosyl transferase, group 1 family protein	
	2	404778	C	T	stop gained	chitinase B
		825824	C	T	stop gained	hypothetical protein
		1045866	C	T	stop gained	hypothetical protein
1270125		C	T	stop gained	potassium uptake protein, TrkH family	
1964935		G	A	stop gained	conserved hypothetical protein	
2019941		G	A	stop gained	transcriptional regulator, TetR family	
2072349		C	T	stop gained	conserved hypothetical protein	
2491611		T	G	stop gained	transporter, EamA family	
2685192		G	A	stop gained	putative membrane protein	
3140849		A	T	stop gained	TPR domain protein	
3542331		G	A	stop gained	DNA-3-methyladenine glycosidase	
3		1211363	C	T	stop gained	putative L-lactate permease
	1689676	C	T	stop gained	putative helicase	

	18645 65	T	C	stop lost and splice region variant	conserved hypothetical protein
	19535 40	A	C	stop gained	transcriptional regulator, GntR family
	22041 32	G	A	stop gained	2,3-dihydroxybenzoate-AMP ligase
	28670 41	G	A	stop gained	phosphoenolpyruvate synthase
	35535 72	G	A	stop gained	putative membrane protein
	50861 44	C	A	stop gained	ABC transporter, ATP-binding protein
4	16624 02	T	A	stop gained	conserved hypothetical protein
5	21054 81	A	T	stop lost and splice region variant	nitroreductase family protein
	25997 35	G	A	stop gained	conserved hypothetical protein
	28155 75	C	T	stop gained	ABC transporter, permease protein
	30578 53	G	A	stop gained	conserved hypothetical protein
	31328 11	G	A	stop gained	putative tryptophanyl-tRNA synthetase
	41067 33	T	C	start lost	conserved hypothetical protein TIGR00486
	47389 83	G	A	stop gained	ABC transporter, ATP-binding protein
6	17486 42	A	T	stop gained	chlorohydrolase family protein
	24238 64	T	C	start lost	hypothetical protein
	26238 95	G	A	stop gained	acetyltransferase, GNAT family
	30832 47	G	A	stop gained	ABC transporter, ATP-binding protein
	47185 00	G	A	start lost	conserved hypothetical protein
3 through 6	15671 21	C	T	stop gained	putative chemotaxis protein MotA
	16624 02	T	A	stop gained	conserved hypothetical protein
	17486 42	A	T	stop gained	chlorohydrolase family protein
	26238 95	G	A	stop gained	acetyltransferase, GNAT family
	47185 00	G	A	start lost	conserved hypothetical protein
4 through 6	15671 21	C	T	stop gained	putative chemotaxis protein MotA
	16624 02	T	A	stop gained	conserved hypothetical protein
	17486 42	A	T	stop gained	chlorohydrolase family protein CDS

	26238	G	A	stop gained	acetyltransferase, GNAT family
	95				
	31328	G	A	stop gained	putative tryptophanyl-tRNA synthetase
	11				
	47185	G	A	start lost	conserved hypothetical protein
	00				
5 through 6	17486	A	T	stop gained	chlorohydrolase family protein
	42				
	24238	T	C	start lost	hypothetical protein
	64				
	25997	G	A	stop gained	conserved hypothetical protein
	35				
	26238	G	A	stop gained	acetyltransferase, GNAT family
	95				
	30832	G	A	stop gained	ABC transporter, ATP-binding protein
	47				
	31328	G	A	stop gained	putative tryptophanyl-tRNA synthetase
	11				
	47185	G	A	start lost	conserved hypothetical protein
	00				

Table S8. Clustered mutations across each population cluster. Positions are relative to the Ames Ancestor reference genome (NCBI: accession AE017334.2)

Population Cluster	SNPs in 2,000bp window	Positions	Gene/Product
1	N=5	5203864 5204158 5205097	sensory box histidine kinase YycG
		5205511 5205579	
	N=4	752178 752186 753164	putative transcriptional activator TenA
		753799	
		"	ABC transporter, ATP-binding protein
		1068476 1068592 1068794	sugar-binding transcriptional regulator, LacI family
	N=3	1069010	
		1260100 1260604 1261317	DNA-binding response regulator
		1262055	
		"	sensor histidine kinase
		1292333 1292469 1292716	stage 0 sporulation regulatory protein
		"	sensor histidine kinase KinD
		1514231 1514852 1515000	acyltransferase family protein
		"	chain length determinant protein
		1587845 1588867 1589053	putative flagellum-specific ATP synthase
		"	unknown
		"	conserved domain protein
		"	conserved hypothetical protein
		3637229 3638405 3638413	1-deoxy-D-xylulose 5-phosphate reductoisomerase
			phosphatidate cytidyltransferase
3643337 3644093 3644570	transcriptional regulator CodY		
"	ATP-dependent hsl protease, ATP-binding subunit hslU		
4552655 4554430 4554446	S-adenosylmethionine synthetase		
"	unknown		
"	phosphoenolpyruvate carboxykinase (ATP)		
4904229 4905132 4905643	carboxyl-terminal protease		
"	cell division ABC transporter, permease protein FtsX		
"	cell division ABC transporter, ATP-binding protein FtsE		
N=2 (in 50bp window)	752178 752186	intergenic	
2	N=4	615306 615982 616129	putative membrane protein
		616443	
	"	putative lipoprotein	
	N=3	656211 656701 657194	sodium/alanine symporter family protein
		1926100 1927629 1927717	cytosolic long-chain acyl-CoA thioester hydrolase family protein
		"	cardiolipin synthetase domain protein
2025641 2026857 2027259	arginyl-tRNA synthetase		

		"	conserved hypothetical protein
		"	hypothetical protein
		"	conserved hypothetical protein
		2180306 2181182 2181540	major facilitator family transporter
		"	putative lipoprotein
		2670535 2671599 2672362	conserved hypothetical protein
		3297984 3298833 3298837	hypothetical protein
		"	conserved hypothetical protein
		"	exonuclease family protein
		3632282 3633195 3634090	DNA polymerase III, alpha subunit, Gram-positive type
		"	hypothetical protein
		"	prolyl-tRNA synthetase
		4799658 4800937 4801446	nucleoside transporter, NupC family
	N=2 (in 50bp window)	264418 264444	intergenic
		339222 339240	intergenic
		2174658 2174696	intergenic
		2631286 2631310	putative glycine oxidase
		4313037 4313083	aquaporin Z
3	N=3	1938186 1938332 1939286	acetyltransferase, GNAT family
		"	D-alanyl-D-alanine carboxypeptidase family protein
		2196206 2196380 2197835	putative exonuclease
		3682161 3682373 3684073	sun protein
		"	methionyl-tRNA formyltransferase
		"	peptide deformylase
		3849500 3850098 3850658	polyphosphate kinase
		3947965 3949024 3949038	DNA-binding protein, CopG family
	N=2 (in 50bp window)	3273849 3273856	intergenic
		3949024 3949038	putative hydrolase, alpha/beta fold family
		4843773 4843811	hypothetical protein
		5072166 5072207	3-hydroxyacyl-CoA dehydrogenase
		5119683 5119709	intergenic
4	None		intergenic
5	N=3	317219 317808 318187	putative membrane protein
		"	putative delta-1-pyrroline-5-carboxylate dehydrogenase
		3140629 3140812 3140945	TPR domain protein
		3347190 3348050 3348749	conserved hypothetical protein
		"	conserved hypothetical protein

			"	putative oligopeptide ABC transporter, oligopeptide-binding protein
		4904797 4905788 4906434	"	cell division ABC transporter, permease protein FtsX
	N=2 (in 50bp window)	1982941 1982986	"	cell division ABC transporter, ATP-binding protein FtsE
		4589569 4589574		respiratory nitrate reductase, alpha subunit
		4751988 4752002		intergenic
6	N=3	1748642 1749441 1750262		chlorohydrolase family protein
			"	conserved hypothetical protein
	N=2 (in 50bp window)	3957722 3957762		6,7-dimethyl-8-ribityllumazine synthase
3-6 combined	N=3	719485 720360 720943		cytochrome aa3 quinol oxidase, subunit IV
			"	cytochrome aa3 quinol oxidase, subunit III
			"	cytochrome aa3 quinol oxidase, subunit I
		1748642 1749441 1750262		chlorohydrolase family protein
			"	conserved hypothetical protein
		1824707 1824833 1824985		conserved hypothetical protein
			"	carboxymuconolactone decarboxylase family protein
		2993945 2994131 2994787		conserved hypothetical protein
			"	LPXTG-motif cell wall anchor domain protein
		3818754 3819811 3820465		GTP-binding protein TypA
			"	hypothetical protein
		3957722 3957762 3959473		6,7-dimethyl-8-ribityllumazine synthase
			"	conserved domain protein
			"	biotin synthase
	N=2 (in 50bp window)	3957722 3957762		6,7-dimethyl-8-ribityllumazine synthase
4-6 combined	N=3	719485 720360 720943		cytochrome aa3 quinol oxidase, subunit IV
			"	cytochrome aa3 quinol oxidase, subunit III
			"	cytochrome aa3 quinol oxidase, subunit I
		1748642 1749441 1750262		chlorohydrolase family protein
			"	conserved hypothetical protein
		1824707 1824833 1824985		conserved hypothetical protein
			"	carboxymuconolactone decarboxylase family protein
		2993945 2994131 2994787		conserved hypothetical protein
			"	LPXTG-motif cell wall anchor domain protein
		3818754 3819811 3820465		GTP-binding protein TypA

			"	hypothetical protein
		3957722 3957762 3959473		6,7-dimethyl-8-ribityllumazine synthase
			"	conserved domain protein
			"	biotin synthase
	N=2 (in 50bp window)	3957722 3957762		6,7-dimethyl-8-ribityllumazine synthase
5-6 combined	N=3	1748642 1749441 1750262		chlorohydrolase family protein
			"	conserved hypothetical protein
		1814297 1814541 1814566		intergenic
		1948662 1949099 1950257		conserved hypothetical protein
			"	acetyltransferase, GNAT family
			"	conserved hypothetical protein
		2423657 2423864 2424305		hypothetical protein
		2993945 2994131 2994787		conserved hypothetical protein
			"	LPXTG-motif cell wall anchor domain protein
		3630983 3631093 3631146		DNA polymerase III, alpha subunit, Gram-positive type
	N=2 (in 50bp window)	1134098 1134120		DNA-binding protein
		1647367 1647389		intergenic
		1814541 1814566		intergenic
		2184985 2185024		mmgE protein
		3957722 3957762		6,7-dimethyl-8-ribityllumazine synthase