



Comparative Genome Analysis of Bacillus sporothermodurans with Its Closest Phylogenetic Neighbor, Bacillus oleronius, and Bacillus cereus and Bacillus subtilis Groups



Rodney Owusu-Darko¹, Mushal Allam², Arshad Ismail², Carlos A. S. Ferreira³, Sílvia D. de Oliveira ^{3,*} and Elna M. Buys ^{1,*}

- 1 Department of Consumer and Food Sciences, University of Pretoria, Private Bag X20, Hatfield, Pretoria 0028, South Africa; rodneyowusu-darko@tuks.co.za
- 2 National Institute for Communicable Diseases, Private Bag X4, Sandringham, Johannesburg 2131, South Africa; mushala@nicd.ac.za (M.A.); arshadi@nicd.ac.za (A.I.)
- 3 Laboratory of Immunology and Microbiology, School of Health and Life Sciences, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre 90619-900, Brazil; cferreira@pucrs.br
- * Correspondence: silviadias@pucrs.br (S.D.d.O.); elna.buys@up.ac.za (E.M.B.)

Received: 14 June 2020; Accepted: 23 July 2020; Published: 4 August 2020



Abstract: Bacillus sporothermodurans currently possesses one of the most highly heat-resistant spores (HRS), which can withstand ultra-high temperature (UHT) processing. Determination of multiple whole genome sequences of *B. sporothermodurans* provided an opportunity to perform the first comparative genome analysis between strains and with B. oleronius, B. cereus, and B. subtilis groups. In this study, five whole genome sequences of *B. sporothermodurans* strains, including those belonging to the HRS clone (SAD and BR12) normally isolated from UHT milk, were compared with the aforementioned *Bacillus* species for gene clusters responsible for heat resistance. In the phylogenomic analysis, B. sporothermodurans, with its closest phylogenetic neighbor, B. oleronius, clustered with B. thermoamylovorans and B. thermotolerans. Heat shock proteins GrpE, GroES, GroEL, and DnaK presented identical sequences for all B. sporothermodurans strains, indicating that differences in functional efficiency are not involved in the thermal resistance variations. However, comparing all species evaluated, B. sporothermodurans exhibited a different gene configuration in the chromosomal region of the heat shock protein GrpE. Furthermore, only B. sporothermodurans strains presented the stage II sporulation protein P gene located in this region. Multisequence alignment and phylogenetic analysis of the ClpB protein showed differences for HRS and non-HRS strains. The study identified ClpC, ClpE, and ClpX as the three ATPases putatively involved in protein disaggregation in B. sporothermodurans. Bacillus sporothermodurans exhibits high homology with other Bacillus species in the DnaK, DnaJ, GroEL, and GroES cluster of genes involved in heat resistance. The data presented here pave the way to select and evaluate the phenotypic effects of genes putatively involved in heat resistance.

Keywords: heat shock; comparative genomics; protein cluster; Bacillus sporothermodurans; Bacillus oleronius; heat resistance

1. Introduction

Comparative genome analysis typically consists of the identification of homologs (orthologues and paralogues), thus elucidating bacterial rearrangements, plasticity, epigenetics, and polymorphisms [1,2]. The afore-mentioned bacterial analysis has especially taken off since the gains in next generation

sequencing (NGS) technologies [3–5]. Recently, NGS technologies evaluating whole genomes have made it possible to study the architecture of bacterial species, with the aim of identifying phylogeny and evolutionary trends [6–8].

Whole genome sequencing (WGS) of particularly industrially important organisms and bacterial pathogens provides critical answers for the biotechnology industry, human health, and wellbeing, as well as the environment. With the advent of bacterial bioprospecting, comparative genomics is the key to identifying new sources of important bacterial metabolites of benefit to the food, agriculture, health, and chemical industries. The study of evolutionary trends and pathogenesis of important disease-causing bacterial species would allow the development of new antibiotics and detailed knowledge of genetic mechanisms, along with the possible use of recombinant technology for pathogen control. Through NGS technologies, significant data are being generated, enabling the understanding of the moderate to high conservation of bacterial species via numerous proteins and unravelling the mechanisms of their adaptation to new niches and competition with other bacteria or forms of life [9].

The genus *Bacillus* comprises low G + C Gram-positive bacteria that are readily found in soil and water and, as such, can easily contaminate the food supply chain. Their ability to form spores in unfavorable conditions enables their survival in harsh environmental conditions that will, otherwise, kill vegetative bacteria. These unfavorable conditions include, but are not restricted to, wet and dry heat, extreme desiccation, ultraviolet and gamma irradiation, and oxidizing agents [10]. *Bacillus* spp. are of importance for the production of various exogenous enzymes such as amylase, protease, lipase, and phytase, used in detergents, starch hydrolysis, textiles, and animal feed applications [11]. *Bacillus* spp. have high growth rates and short cycles in industrial fermentations, making them good candidates for extracellular enzyme production [9]. Several species in this genus can ferment in acid, neutral, and alkaline pH ranges and are engineered to produce nucleotides, riboflavin, and ribose, amongst a variety of other industrial products. Industrially important *Bacilli* include *B. subtilis*, *B. licheniformis*, *B. amyloliquefaciens*, and *B. megaterium*. On the other hand, a few species, such as *B. cereus* and *B. anthrax*, infect humans, causing food-borne illness and anthrax, respectively.

Bacillus sporothermodurans is an emerging highly heat-resistant spore (HRS)-forming bacteria of interest to the dairy industry. HRS can survive in foods treated with ultra-high temperature (UHT), affecting quality and causing significant economic losses. After contamination of UHT milk products, these spores can germinate and grow in the stored milk, possibly reaching 10⁵ colony-forming unit (CFU)/mL. Taking into account that UHT milk products are considered commercially sterile when there is a total count of ≤ 10 CFU/0.1 mL, contamination with *B. sporothermodurans* may result in exceeding this limit. Even though *B. sporothermodurans* neither affects the pH nor the sensory quality and rarely causes any characteristic spoilage, its presence contravenes good manufacturing practice [12], because of the non-sterility of UHT milk products. Although no known pathogenic properties have been characterized to date, it is important to prevent its occurrence in UHT milk for quality purposes. Additionally, molecular approaches are needed to properly identify *B. sporothermodurans* and to differentiate strains isolated from UHT milk with likely higher heat resistance and those isolated from other sources including raw milk, milking equipment and feed concentrate [13]. Differences in heat resistance are thought to be clonal, which explains the growing interest in the identification of evolutionary associations and the understanding of its high heat resistance. In this study, the whole genomes of B. sporothermodurans strains isolated from UHT milk from South Africa and Brazil, the type strain, and *B. oleronius*, its closest phylogenetic neighbor, were sequenced [14,15], with the aim of comparing proteins involved in heat resistance between B. sporothermodurans and other Bacillus species to better understand this phenotype in *B. sporothermodurans*. By using molecular comparisons, we report various protein family clusters shared with other *Bacillus* species. To the best of our knowledge, this is the first study comparing B. sporothermodurans genomes with those of other Bacillus species.

2. Materials and Methods

2.1. Bacterial Isolation and Identification

Bacillus sporothermodurans strains SAD and SA01 were isolated from UHT milk sourced from local producers in South Africa, and *B. sporothermodurans* strain BR12 was isolated from UHT milk produced in Brazil. The type strains *B. sporothermodurans* DSM 10599 and *B. oleronius* DSM 9356 were obtained from the Leibniz Institute DSMZ—German Collection of Microorganisms and Cell Cultures GmbH (DSMZ). Growth of *B. sporothermodurans* isolates was performed on brain heart infusion agar (Hampshire, Oxoid, UK) and incubated at 37 °C for 48 h. *Bacillus oleronius* was grown on nutrient agar (Oxoid, UK) at 37 °C for 24 h. Genomic DNA was extracted using the ZR Bacterial DNA Miniprep kit (Zymo Research, Irvine, CA, USA). The DNA was quantified using the Qubit instrument and the dsDNA BR Assay kit (Life Technologies, Grand Island, NY, USA). Molecular confirmation of *B. sporothermodurans* strains and SA01 and BR12 as HRS was performed by PCR, as previously described [13,14].

2.2. Genome Sequencing and Analysis

Bacillus sporothermodurans strains SAD, SA01, BR12, and DSM 10599, as well as *B. oleronius* DSM 9356, were sequenced and annotated as previously described [15,16]. Subsequent analyses in this study are based on the strains mentioned above. *Bacillus sporothermodurans* B4102 and the selected *B. subtilis* and *B. cereus* group members were sourced from NCBI. Subsystem-based annotation was undertaken for all 12 *Bacillus* species, employing Rapid Annotation using Subsystem Technology (RAST) (Chicago, IL, USA) [17–19].

Genome sequences for *B. sporothermodurans* strains SAD, SA01, BR12, DSM 10599, and B4102 (accession numbers NAZD01000000, NAZB01000000, NAZA01000000, NAZC01000000, and LQYN00000000, respectively) and *B. oleronius* DSM 9356 (MTLA01000000) were used in this study. The *B. subtilis* group members included were *B. subtilis* subsp. *subtilis* str. 168 (CP019663), *B. licheniformis* DSM 13 = ATCC 14580 (CP000002), and *B. amyloliquefaciens* DSM 7 (FN597644). The *B. cereus* group members included *B. cereus* ATCC 10987 (CP026375), *B. anthracis* str. 'Ames Ancestor' (AE017336), and *B. thuringiensis* strain ATCC 10792 (CP020754).

Bioinformatics tools, including RAST and the Pathosystems Resource Integration Centre platform (PATRIC) (Chicago, IL, USA), were used to identify genes and protein families [17,20] of *B. sporothermodurans* strains, *B. oleronius* DSM 9356, and selected *B. subtilis* and *B. cereus* group members. Molecular Evolutionary Genetics Analysis (MEGA) v. 7.0 (University Park, PA, USA) was used for phylogenetic analysis [20–22].

A phylogenetic tree was constructed by comparing the conserved protein sequences obtained in this study with sequences of related *Bacillus* spp. selected from NCBI, using the programs RAxML (Randomized Axelerated Maximum Likelihood) (Heidelberg, Germany) [23] and FigTree v. 1.4.3 (Edinburgh, UK) [24]. The Phyre2 suite of tools was used to predict the tertiary structure of protein GrpE [25] at 100% model confidence for all four protein structures. Protein sequence alignment was undertaken with BioEdit v. 7.2.1. (Raleigh, NC, USA) [26]. The Jensen–Shannon divergence approach was employed to analyze the conservation of amino acid positions among the strains [27], and amino acid conservation was based on a relative amino acid conservation score of 75%. Single nucleotide polymorphisms (SNPs), based on phylogeny among the strains, were identified using CSI Phylogeny server v. 1.4 (Lyngby, Denmark) [28]. The WGS was aligned to the reference DSM 10599, and subsequently SNPs were called and filtered. Based on CSI phylogeny, we validated WGS reads and inferred relationships based on the concatenated alignment of SNPs.

3. Results

3.1. Wide Variations in Bacillus Species

To compare *B. sporothermodurans* strains (SAD, SA01, BR12, DSM 10599, B4102) with other closely related *Bacillus* species, the genomic features of *B. sporothermodurans* strains, as determined and deposited in GenBank, were compared to type strains from the *B. subtilis* group (*B. subtilis*, *B. licheniformis*, *B. amyloliquefaciens*) and the *B. cereus* group (*B. cereus*, *B. anthracis*, *B. thuringiensis*). The average genome size of the *B. sporothermodurans* strains was 3.7 Mb, on average 25% smaller than that of *B. oleronius*, 10% smaller than that of the *B. subtilis* group strains, and 30% smaller than that of the *B. cereus* group strains (35.4%), but lower than that of the *B. subtilis* group strains (45.46%). The presence of plasmids was detected only for the *B. cereus* group strains. Concerning antimicrobial resistance genes, 33 were detected in *B. sporothermodurans*, 51 in *B. subtilis*, and 46 in *B. cereus* (Figure 1).



Figure 1. *B. sporothermodurans* SA01 (**A**), *B. oleronius* (**B**), *B. subtilis* subsp. *subtilis* str. 168 (**C**), and *B. cereus* ATCC 10987 (**D**) possess unique genomic regions in comparison with each other. The outer circle designates the genome's coordinates in mega base pairs (Mbp). The blue circle denotes the number of contigs, and the green and purple circles represent the coding sequences (CDS), forward and reverse, respectively, with white spaces between the CDS accounting for hypothetical proteins. Non-CDS features are shown by the light blue circle. The proceeding red and orange circles represent regions encoding antimicrobial determinants and virulence factors, respectively. The two inner black circles depict the G + C contents and the skew.

3.2. Phylogenetic Analysis Depicts Distinct Clusters of B. sporothermodurans Separate from the B. subtilis Group

Figure 2 shows the phylogenetic relationship of *Bacillus* species. *Bacillus oleronius* showed higher relatedness to *B. sporothermodurans* strains and both *B. thermotolerans* and *B. thermoamylovorans,* two highly heat-resistant bacteria grouped in the *B. sporothermodurans* cluster (cluster A). Other clusters included the *B. subtilis* group (cluster B) and the *B. cereus* group (cluster C). Cluster D mostly consisted of alkaliphiles, including *B. akibai* and *B. hemicellulosilyticus*, used in the degradation of cellulose and starch in the biotechnology industry.



Figure 2. Phylogenetic analysis of *Bacillus* species showed distinct clusters for *B. sporothermodurans* and its group (**A**), *B. subtilis* group (**B**), *B. cereus* group (**C**), and selected alkaliphiles from *Bacillus* genus (**D**). The tree was constructed using the Randomized Axelerated Maximum Likelihood (RAxML) and FigTree programs.

3.3. SNP Analysis

The phylogenetic tree constructed from SNP calling (Figure 3) shows the two *B. sporothermodurans* strains belonging to the HRS clones clustered together. The number of SNPs between those strains was 183, as opposed to 11,025 SNPs for the two most distant strains.



Figure 3. Single nucleotide polymorphism (SNP) analysis showing the phylogeny of the five *B. sporothermodurans* strains SAD, SA01, BR12, DSM 10599, and B4102, with type strain DSM 10599 as reference. The phylogenetic tree was derived from the Newick files of CSI phylogeny and processed with FigTree.

3.4. Protein Clusters Involved in Heat Resistance

The protein families identified as responsible for heat resistance are shown in Figure 4, including a total of 21,422 protein clusters, with a core of 482 protein clusters common to all 12 *Bacillus* species. *Bacillus sporothermodurans, B. oleronius, B. subtilis* group, and *B. cereus* group presented 679, 2559, 659, and 2476 clusters, respectively. *Bacillus sporothermodurans* had a total of 38, 68, and 431 clusters common to the *B. subtilis* group, *B. cereus* group, and *B. oleronius,* respectively. We highlight GrpE since it exhibited less homology between *B. sporothermodurans* strains and other *Bacillus*, as indicated by multiple sequence alignments. Figure 5 depicts the tertiary protein structures of GrpE from *B. sporothermodurans* and from its closest phylogenetic neighbor, *B. oleronius*, in comparison with *B. cereus* and *B. subtilis*. The protein structure of the four *Bacillus* species had, on average, a 64% identity with the GroES family of proteins and a 37.5% average identity with the human mitochondrial chaperonin symmetrical football complex.

Figure 6 shows the nucleotide and amino acid sequence alignments of the heat shock protein GrpE, encoded by the *dna*K operon. The protein sequence alignment depicts areas of conservation and mutations in the Bacillus species under study, and no differences were observed among the B. sporothermodurans strains. Although GrpE, DnaK, and DnaJ are present in similar positions in all species, the hypothetical radical SAM family enzyme (gene 5) swaps locations with the ribosomal RNA small subunit methyltransferase (gene 7) only in *B. sporothermodurans*. Additionally, the ribosomal RNA small subunit methyltransferase in B. sporothermodurans consists of a 251-amino acid chain, as compared to 379 amino acids in B. oleronius and B. cereus and 367 in B. subtilis. The amino acid chain length for the hypothetical radical SAM family enzyme and ribosomal protein L11 methyltransferase in B. sporothermodurans was 379 and 313, respectively, as compared to 313 and 251, respectively, in B. oleronius, 312 and 257, respectively, in B. subtilis, and 313 and 250, respectively, in B. cereus. The *dna*K gene cluster is highly conserved in the *B. sporothermodurans* strains and exhibited identical amino acid length and configuration. There were no differences in the *grpE* position within the *dna*K gene cluster with respect to the *B. sporothermodurans* strains. Conserved genes were given the same number and are shown with a grey background color. The genomic arrangement of grpE, as well as associated genes responsible for heat shock in the *dna*K gene cluster, are shown in Figure 7. Sequence alignments and phylogenetic analysis undertaken for chaperon protein ClpB and ATPases ClpC, ClpE and ClpX showed variable regions. Regarding ClpB, there were nine variable regions out of the 862 amino acids in the *B. sporothermodurans* strains, with BR12 and SA01 sharing the same sequence,

although SA01 presented two additional amino acids. The ClpC, ClpE, and ClpX exhibited one variable region each out of 813, 708, and 422 amino acids, respectively. Phylogenetic analysis using amino acid sequences for ClpB, ClpC, ClpE, and ClpX are shown in Figure 8.



Figure 4. Heat map of genomes of *Bacillus* species and associated protein clusters responsible for heat resistance. The heat map was inferred using the PATRIC protein families' utility and clustered by genomes using Pearson's correlation and pairwise average linkage. Cell color represents the number of proteins from a specific genome in each protein family.



Figure 5. Comparison of predicted tertiary structure of the heat shock protein GrpE as identified in *B. sporothermodurans* strains SA01, SAD, BR12, DSM 10599, and B4102 (**A**), *B. oleronius* DSM 9356 (**B**), *B. cereus* ATCC 10987 (**C**), and *B. subtilis* subsp. *subtilis* str. 168 (**D**). Models were predicted using the Phyre2 suite of tools and at 100% model confidence for all four protein structures. Protein sequence alignment was performed with BioEdit v. 7.2.1.

(A)

B. sporothermodurans str. SA01	ATGGATAAAO	G AAGCAAATGA	AAAATATAGA	GCGCAAAATC	TTGCGACCAA	TTTATTACCG	GCCATTGATA	ATTTAGAACG CGCAATGCAA
B. sporothermodurans str. DSM 10599 B. sporothermodurans str. SAD	•••••			•••••			•••••	•••••
B. sporothermodurans str. B4102								
B. sporothermodurans str. BR12 B. oleronius DSM 9356	CCGT	.GG	AG		.A.TTA			CA. AT
B. anthracis str. 'Ames Ancestor'	G	C .GTGC		AG	.A.TTT.AG.	.A.TGA	TC	TA. A
B. cereus ATCC 10987 B. thuringiensis str. ATCC 10792	G	C .GTGC C .GTGC		AG AG	TTT.GG.	.A.TGA CA.CGA	TC	TA. AG TA. A
B. subtilis subsp. subtilis str. 168	G.TG.	GTCCC.	CC.T	T.TA	.C.TTG.	GC.G	TCC.	GT AGC.TG
B. licheniformis DSM 13 = ATCC 14580	CTC.		GCC.C	T.CGCGCA	.CATC.G.G.	GC.G		.CTGC.TG

....|...|||||||||||||||||| 100 110 120 130 140 150 160 170 180

	100	, 11	0 12	0 13	0 14	0 15	0 100) 1/(J 19
B. sporothermodurans str. SA01	ATAACACCTG	AAAATGAACA	AACAAAACAA	TTGCTACAAG	GTGTTGAGAT	GGTATACCGT	AATATTATCG	ATGCATTGAA	GCAAGAAGGA
B. sporothermodurans str. DSM 10599									
B. sporothermodurans str. SAD									
B. sporothermodurans str. B4102									
B. sporothermodurans str. BR12									
B. oleronius DSM 9356		GG	G G	A G	а а	т д	тт	c	а т
B. anthracis str. 'Ames Ancestor'	G GGA G AA	CTG G	G TCC	СТ	AGA	GC T	CATGCA	a a	та т
B. cereus ATCC 10987	C CA C AA	CTC C	TC CTCC	5 5T	CA C A	тт	C AT CT A	A CA C	55 G T
B thuringiensis str ATCC 10792	C CCA C CA	CTC C			. GA. G A	CC 7	C AT CT A	A. CC C	TA
R subtilis subsp. subtilis str. 168	C. TCA. C. C	c					C. CC. CC. A	.AGCC	77
P. amulalizuofacione DSM 7	G.16A.G.C.		GGAGT		.AA.GA		C.GC.CG.A.	.A	AA
B. amyioliquejaciens DSIVI 7	GGA.G	GCG	GAGC	· · · · · G · · · ·	A.GA	CCGC	C.GC.GA.	.CCC.TG.	AAC
B. licheniformis DSM 13 = ATCC 14580	CGAT	.CG	GGAGC	СТ	.AA.GA	GC	C.AT.G.	.GTC.T	AA.T

B. sporothermodurans str. SA01 B. sporothermodurans str. DSM 10599 B. sporothermodurans str. SAD B. sporothermodurans str. B4102 B. sporothermodurans str. BR12 B. oleronius DSM 9356 G...G. ... A. T. T. T. C. G. ... T. GAA. C. ... TG. CAGTGAA. T. AA. A. G. A. . GC. ... A. T. T. T. C. G. ... TC ... GAA. C. ... G. ... TA ... CAGCGAA. T. AA. A. B. anthracis str. 'Ames Ancestor' B. cereus ATCC 10987 B. thuringiensis str. ATCC 10792 ${\tt G.G...G.} \ldots {\tt A..T.} {\tt T..T..C.G} \ldots {\tt T...GAA..C.} \ldots {\tt TG....TG...AGTGAA..T.AA..A..C}$ B. subtilis subsp. subtilis str. 168 B. amyloliquefaciens DSM 7 B. licheniformis DSM 13 = ATCC 14580 G.....CA....CCAT.A.. C...G.GC......CA.G.C.....CA.G.C.....G.T...G.AGCG.A ...AA..A..C

|....|
|...|
|....|
|....|
|....|

 280
 290
 300
 310
 320
 330
 340
 350

	280	. 29	0 300	310	0 320	0 330) 340		0
B. sporothermodurans str. SA01	ATTGTTATCG	AAGAGTTCCA	AAAAGGTTAT	ATTTTGAAGG	ATCGTGTTAT	TCGTCCTTCA	ATGGTGAAAG	TGAATCAATA	P
B. sporothermodurans str. DSM 10599									
B. sporothermodurans str. SAD									
B. sporothermodurans str. B4102									
B. sporothermodurans str. BR12									
B. oleronius DSM 9356	C	AT		A	A	T	AG.	.AC	
B. anthracis str. 'Ames Ancestor'	GCGAG.T.	A		.AAC.AA.	.CG	CA	A	.A	
B. cereus ATCC 10987	GCGG.G.	A		.AAAA.	.CA	CAG	A	.A	
B. thuringiensis str. ATCC 10792	GCGG.G.	A		.AAAA.	.CG	CA	A	.A	
B. subtilis subsp. subtilis str. 168	G.T.	.GAA.G	C	.AGC	cc.	cc	C		
B. amyloliquefaciens DSM 7		AC.G	CC	.AAC	C	С	A		
B. licheniformis DSM 13 = ATCC 14580	GCACG.T.	.GAC.G	C	.AAC.T	.CGG	СТ	A	.C	

Figure 6. Cont.

		1	0 20) 30	0 40) 50	0 60) 70
	B. amyloliquefaciens DSM 7	MSEEKQTAEQ	VEAAEQ-EEM	TEQAASEEQH	EETVG	QEEDLQHQID	ELQGLLDEKE	NKLLRVQADF
	B. sporothermodurans str. B4102	MAEEIKQ	ENQQEQIFAE	E-ETNSDEQV	DSVETEKN	EKNPEDEIIK	ELQAKLEESE	NRYLRLRADF
	B. sporothermodurans str. SA01	MAEEIKQ	ENQQEQIFAE	E-ETNSDEQV	DSVETEKN	EKNPEDEIIK	ELQAKLEESE	NRYLRLRADF
	B. sporothermodurans str. SAD	MAEEIKO	ENOOEOIFAE	E-ETNSDEOV	DSVETEKN	EKNPEDEIIK	ELOAKLEESE	NRYLRLRADF
	B sporothermodurans str BR12	MAEEIKO	ENCOFOTFAE	E-ETNSDEOV	DSVETEKN	EKNPEDETIK	ELOAKLEESE	NRYLRLRADE
	B. sporothermodurans str. DSM 10599	MAEEIKO	ENOOEOIFAE	E-ETNSDEOV	DSVETEKN	EKNPEDETIK	ELOAKLEESE	NRYLRLRADE
	B oleronius DSM 9356	MAEEKNOEMT	EETVETIFEE	ETEVNSSEOT	ETKEKEES	DANDATE	ELOKKLDESE	NRYLELEADE
	B licheniformis DSM 13 = ATCC 14580	MAEEKONEEL	NEOFELNETE	AFTAFAFOTA	AFADAPAFET	OTEMLEKOLK	ELOERLEEKE	NKLLRVOADE
	B anthracis str 'Ames Ancestor'	VEEDNEOUVE	EVKEDO	VEEAUTDENS	FETUEEKS	FADITOFKUD	FLOAKLTETE	GPTIPLOADE
	B subtilis subsp subtilis str 168	MSEEKOTVEO	NETERO-ETT	FEORADEOO	FFTN	ESELLONOIN	FLOCITEEKE	NKLIDVOADE
	B. corous ATCC 10997	TREENCOUR	MELEEX-EIT	TREVINGEDO	EETN EKO	ESERPTINE	ELQONIMEME	CDMIDIOADE
	B. thuringionsis str. ATCC 10702	VEEKNEQVVE	FALFFALFAC	VEEAVISEDS	FFIAFFV2	FWATPŐFKAD	FTÓWUTIFIE	GRMLKLQADF
	B. thuringiensis sti. AICC 10792	and and this time and any and and and and	the are not not use the fact and the	the same and and the same and and and	NOT THE ME OF THE ME OF THE DOT THE	nes and and and the lost and and the set	the same and the same same and the same	and she one and one and and and and and
		••••			• • • • • • • •			
		8	0 90	100	0 110) 120	0 130	140
	B. amyloliquefaciens DSM 7	ENYKRRSRLE	MEAAQKYRSQ	NVVTEILPAL	DNFERALQVE	AESEQTKSLL	QGMEMVRRQL	IDALEKEGVE
	B. sporothermodurans str. B4102	DNFRRRVQID	KEANEKYRAQ	NLATNLLPAI	DNLERAMQIT	PENEQTKQLL	QGVEMVYRNI	IDALKQEGIE
	B. sporothermodurans str. SA01	DNFRRRVQID	KEANEKYRAQ	NLATNLLPAI	DNLERAMQIT	PENEQTKQLL	QGVEMVYRNI	IDALKQEGIE
	B. sporothermodurans str. SAD	DNFRRRVQID	KEANEKYRAQ	NLATNLLPAI	DNLERAMQIT	PENEQTKQLL	QGVEMVYRNI	IDALKQEGIE
(D)	B. sporothermodurans str. BR12	DNFRRRVQID	KEANEKYRAQ	NLATNLLPAI	DNLERAMQIT	PENEQTKQLL	QGVEMVYRNI	IDALKQEGIE
(D)	B. sporothermodurans str. DSM 10599	DNFRRRVQID	KEANEKYRAQ	NLATNLLPAI	DNLERAMOIT	PENEQTKOLL	QGVEMVYRNI	IDALKQEGIE
	B. oleronius DSM 9356	DNFRRRVOID	REANEKYKAO	SLVTNLLPAI	DNFERAMOIT	PEDEOAKOLM	OGVEMVYRNI	IDALKOEGVE
	B. licheniformis DSM 13 = ATCC 14580	ENYKRRARLD	LEAAEKYRSO	RIISDLLPAL	DNFERALOID	PDNEOTKSLL	OGMEMVHRO1	LEALKNEGVE
	B. anthracis str. 'Ames Ancestor'	ENYKRRVOMD	KOAAEKYRAO	SLVSDILPAL	DNFERAMOVE	ATDEOTKSLL	OGMEMVHROL	LEALNKEGVE
	B. subtilis subsp. subtilis str. 168	ENYKERSELE	MEASOKYRSO	NTVTDLLPAL	DSFERALOVE	ADNEOTKSLL	OGMEMVHROT.	VEALKKEGVE
	B. cereus ATCC 10987	ENYKERVOMD	KODDEKYRDO	SLVSDILPAL	DNEERAMOVE	ATDEOMKSLI.	OGMEMVYROI.	LEAMTKEGVE
	B thuringiensis str. ATCC 10792	======MD	KODDEKABDO	SLUSDILPAL	DNFFRAMOUF	ATDEOTKSLI.	OGMEMUHPOL	LEALTKEGVE
	D. than ingle is but in the 20102	MD	NYAADINI NAY	ODVODIDEAD	DIAL BUHHŐAR	AIDBQ1105DD	200021VIII/2D	DEADINEOVE
	Den de l'anné de la constant	15	0 T90	1/1	0 18(190	J	
	B. amyloliquefaciens DSM 7	AIEAVGQEFD	PNLHQAVMQV	EDENFGSNIV	IEELQKGYKL	KDRVIRPSMV	KVNQ*-	
	B. sporothermodurans str. B4102	MIEAVGKEFD	PNFHQAIMQG	EDESFGSNIV	IEEFQKGYIL	KDRVIRPSMV	KVNQ*-	
	B. sporothermodurans str. SA01	MIEAVGKEFD	PNFHQAIMQG	EDESFGSNIV	IEEFQKGYIL	KDRVIRPSMV	KVNQ*-	
	B. sporothermodurans str. SAD	MIEAVGKEFD	PNFHQAIMQG	EDESFGSNIV	IEEFQKGYIL	KDRVIRPSMV	KVNQ*-	
	B. sporothermodurans str. BR12	MIEAVGKEFD	PNFHQAIMQG	EDESFGSNIV	IEEFQKGYIL	KDRVIRPSMV	KVNQ*-	
	B. sporothermodurans str. DSM 10599	MIEAVGKEFD	PNFHQAIMQG	EDESFGSNIV	IEEFQKGYIL	KDRVIRPSMV	KVNQ*-	
	B. oleronius DSM 9356	VIEAVGKEFD	PNFHQAIMQG	EDENFDSNIV	IEEFQKGYIL	KDRVIRPSMV	KVNQ*-	
	B. licheniformis DSM 13 = ATCC 14580	QIPSVGEQFD	PNMHQAVMQV	EDEAYESNAV	VEELQKGYKL	KDRVIRPSMV	KVNQ-*	
	B. anthracis str. 'Ames Ancestor'	VIEAVGKQFD	PNEHQAIMQV	EDSEFESNAV	VEEFQKGYKL	KDRVIRPSMV	KVNQ*-	
	B. subtilis subsp. subtilis str. 168	AIEAVGQEFD	PNLHQAVMQA	EDENYGSNIV	VEEMQKGYKL	KDRVIRPSMV	KVNQ*-	
	B. cereus ATCC 10987	AIEAVGKOFD	PHEHQAVMOV	EDSEFESNAV	VEEFQKGYKL	KDRVIRPSMV	KVNQ*-	
	B. thuringiensis str. ATCC 10792	VIEAVGKOFD	PNEHOAIMOV	EDSEFESNAV	VEEFOKGYKL	KDRVIRPSMV	KVNO*-	

Figure 6. Nucleotide sequence alignment of the heat shock protein, GrpE (**A**), showing conserved regions, and the sequence alignment of its corresponding amino acid sequence (**B**), indicating the effects of the various mutations of the nucleotides on the protein sequence. Sequences were aligned with BioEdit v. 7. (**A**,**B**) were divided into four and three sections, respectively, to aid viewing.



Figure 7. Comparison of chromosomal regions of the heat shock protein GrpE (1; red) of *B. sporothermodurans* SA01, *B. oleronius* DSM 9356, *B. subtilis* subsp. *subtilis* str. 168, and *B. cereus* ATCC 10987. Sets of genes with similar sequences are grouped with the same number and color. Genes whose relative positions are conserved in at least four other species are functionally coupled and share grey background boxes. Proteins of interest in the region are as follows: 2, heat-inducible transcription repressor HrcA; 3, chaperon protein DnaK; 4, chaperon protein DnaJ; 5, hypothetical radical SAM family enzyme in heat shock genes cluster; 6, ribosomal protein L11 methyltransferase; 7, ribosomal RNA small subunit methyltransferase; 8, translation elongation factor LepA; 9, tRNA-t(6)A37 methylthiotransferase; 10, endopeptidase spore protease Gpr; 11, SSU ribosomal protein S20p; 12, stage II sporulation protein P (SpoIIP); 14, sodium-dependent phosphate transporter; 15, transamidase GatB domain protein; 20, transcriptional regulator, HxIR family; 22, hypothetical protein; **y**, *yqeP/yqx*A gene.



Figure 8. Molecular phylogenetic analysis by the maximum likelihood method of chaperon protein ClpB (**A**) and ATPases ClpC (**B**), ClpE (**C**), and ClpX (**D**). The evolutionary history was inferred by

using the maximum likelihood method based on the Jones-Taylor-Thornton (JTT) matrix-based model [29]. The bootstrap consensus tree inferred from 1000 replicates was used to represent the evolutionary history of the taxa analyzed [30]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, followed by selecting the topology with superior log likelihood value. The analysis involved five amino acid sequences. All positions containing gaps and missing data were eliminated. There were 862, 813, 708, and 422 positions for ClpB, ClpC, ClpE, and ClpX, respectively, in the final dataset. Evolutionary analyses were conducted in MEGA7 [22].

4. Discussion

Bacillus sporothermodurans is an emerging highly heat-resistant spore-forming bacterial species that can significantly affect the quality of UHT and other heat-processed foods. Despite its increasing importance, limited information is available on its genetic characteristics and, most importantly, the relatedness to other *Bacillus* species of importance, especially regarding the food industry.

Bacillus sporothermodurans strains clustered with *B. oleronius* and other closely related *Bacillus* species, such as *B. lentus* and *B. firmus* (Figure 2, cluster A), which were initially considered as their closest relatives [31] through 16S rDNA partial sequencing. Whole genome phylogenetic analysis suggests that *B. firmus* is more distant from *B. sporothermodurans*, whereas *B. fordii* is a closer relative. *Bacillus thermoamylovorans*, another emerging highly heat-resistant *Bacillus* species isolated from milk and known to be almost as heat-resistant as *B. sporothermodurans* [32], also grouped within the *B. sporothermodurans* cluster. Similarly, the moderately heat-resistant *B. thermotolerans*, initially thought to be closely related to *B. firmus* [33], was included in cluster A, despite recent calls for it to be given a new genus, *Quasibacillus thermotolerans* [34]. The relatedness of the whole genome sequence of *B. thermotolerans* with *B. coagulans*, *B. firmus*, and *B. lentus*, as evident in the present study, supports its maintenance in the genus *Bacillus*. The categorization of bacterial populations is facilitated by using sequence similarity thresholds without being necessarily biologically relevant [35].

High heat-resistance in *B. sporothermodurans* has been attributed to the differences among strains [14,36]. Consequently, SNP evaluation, in addition to the pangenome, has been advocated for its use in tracking clonal strains. This combination of resources may help to remove the stumbling block of the high number of genes with unknown function and to provide a super-resolution view into bacterial subpopulations [7]. Given the proposed function of the pangenome and SNPs in adoptive evolution, the latter may be employed as a predictive genotypic marker to identify characteristics, such as high heat-resistance, attributed to strain differences, offering much needed insight into niche specification and adaptation [37,38]. The SNPs between the reference strain (DSM 10599) and the two strains belonging to the HRS clone, SA01 and BR12, were of 97 and 112, respectively, and 183 between BR12 and SA01. Although strain DSM 10599 is not characterized as an HRS clone feature by PCR, it is genetically closer to the two strains belonging to the HRS clone. This is because bacterial species may rely on a relatively small number of unique regions to generate variation [39]. The HRS South African SA01 strain clustered with the Brazilian BR12 strain, despite the geographic distance and apparently unconnected milk supply and production chains. Higher numbers of SNPs were found in strain B4102 (isolated from Indian curry) as opposed to strains SA01, SAD, BR12, and DSM 10599 (isolated from UHT milk), indicating a possible influence of niche adaptation on DNA sequence variation.

To maintain cellular proteostasis, cells need various chaperone pathways that enhance inherent protein folding [40]. These chaperones target misfolded, unfolded, and aggregated polypeptides for reactivation or degradation [41–43]. The misfolding of proteins may be a consequence of the effects of environmental stresses, such as increased heat exposure [40]. The chaperone pathways are especially important in organisms such as *B. sporothermodurans* and other *Bacillus* species exposed to increased heat conditions during food processing. The presence of disulfide bonds plays a critical role in the structural stabilization of intracellular proteins, with thermophiles normally rich in these bonds [44,45].

The response to heat shock, including such extreme conditions as thermal food processing, is mediated by heat shock proteins (HSPs). In addition to heat shock, HSPs are induced by other physiological stresses, such as exposure to cold, ultraviolet light, or during cell healing [46]. The main HSPs involved in prokaryotes are GroES, GrpE, DnaJ, GroEL, DnaK, HtpG, ClpB, ClpA, and ClpX [41,46–49]. Excluding ClpA, all the main heat shock proteins listed above are present in B. sporothermodurans and with identical amino acid sequences barring ClpB, in which the two HRS clones (SA01 and BR12) showed a high identity. The molecular chaperones above support forward protein folding, thus preventing aggregation. They are unable to facilitate protein disaggregation as aggregates form [50]. However, ClpB has the remarkable ability to rescue stress-damaged aggregated proteins [50]. This is achieved by the extraction of polypeptides from aggregates through forced unfolding and translocation via the ClpB central cavity and by the subsequent release for chaperone-mediated refolding [51]. The difference in the sequence of ClpB from the two HRS strains deserves further analyses to evaluate their possible influences on the highly heat-resistant phenotype. In fact, HSP100/Clp proteins are ATP-dependent chaperones that transfer proteins into the chamber of an associated barrel-like protease complex (ClpP, ClpQ) [52]. These ATPases, including ClpA, ClpC, ClpE, ClpX, and ClpY, deliver partially unfolded substrates to ClpP and ClpQ, the corresponding proteolytic component [52,53]. This process is crucial to highly heat-resistant Bacillus spp. during periods of stress and sporulation [54]. Previous studies identified four Clp ATPases (ClpC, ClpE, ClpX, and ClpY) [52,54] in the B. subtilis group. In this study, the coding sequence for ClpA was identified in addition to the previous Clp ATPases for B. subtilis, but not for B. sporothermodurans strains. The B. cereus group used had an identical complement of ATPases in relation to the B. subtilis group. Bacillus sporothermodurans and B. oleronius exhibited coding sequences for ClpC, ClpE, and ClpX. Furthermore, ClpB was identified in B. subtilis, *B. sporothermodurans*, and *B. oleronius*. The ClpB is not involved with proteases and rather works with the chaperones DnaJ, DnaK, and GrpE to repair heat-induced protein damage [55]. The ClpX targets specific proteins for degradation directly or with substrate-specific adaptor proteins. Thus, the ClpB, ClpA, ClpX protein complex facilitates the tolerance to extreme temperatures in various prokaryotes by remodeling and degrading aggregated proteins [40]. Specifically, the ClpX and ClpP protein group plays a role in cell viability at heat stress under conditions of limiting levels of the DnaK system [53].

Apart from HSP20 (present only in the *B. cereus* group), heat-inducible transcription repressor HrcA (detected in B. subtilis and B. sporothermodurans) and DnaK suppressor protein (detected in B. amyloliquefaciens and B. subtilis), the remainder of the genes of the dnaK cluster, had homologs in all Bacillus groups evaluated. The hrcA is the first cistron of the B. subtilis dnaK operon and encodes a negative regulator of class I heat shock genes (*dna*K and *gro*E operons) [46,56]. The DnaK, DnaJ, and GrpE form a cellular chaperone machinery involved in the repair of heat-induced protein damage [48]. Many of these proteins shared amongst the Bacillus species are homologs of each other, with differences along the amino acid chain. In this sense, Figure 6 depicts nucleotide and amino acid alignment of GrpE, a chaperone part of the *dna*K operon that, in conjunction with GroEL (a 60-kDa family chaperone), is responsible for the renaturation of heat-denatured proteins [48]. Differences in the arrangement of genes in the chromosomal region of grpE among the Bacillus species studied (Figure 7) are evident, but, in fact, the effects on the heat resistance or other biological functions of the change between the hypothetical radical SAM family enzyme and the ribosomal RNA small subunit methyltransferase in *B. sporothermodurans* are not known. The hypothetical radical SAM family enzyme may modulate the action of *hrcA* in *B. sporothermodurans*. Indeed, it has been shown that HrcA acts as a thermo-sensor in *B. subtilis* and *B. thermoglucosidasius*, the latter possessing significantly higher heat resistance than the former [57]. Another difference is the presence of a stage II sporulation protein P gene (SpoIIP; region 12 of Figure 7) only in B. sporothermodurans, placed just upstream of yqeP. The SpoIIP may be involved in heat resistance enhancement, as it has been identified as an autolysin with peptidoglycan hydrolase activity; it also participates in the engulfment process during sporulation [58]. These differences may well involve transcription regulation and protein protection mechanisms, resulting in differences in resistance to stresses, such as heat shock.

Besides the sequence variation in the chromosomal region of *grpE*, the predicted tertiary protein structure of GrpE from the *Bacillus* species studied showed differences. The GrpE structure within *B. sporothermodurans* strains (HRS and non-HRS) showed no differences and, as such, cannot be implied in the difference in high heat resistance between these strains. In addition, GrpE from *B. sporothermodurans* and *B. oleronius* presented almost identical amino acid chains, only differing on the 16th amino acid, where the former's alanine is replaced by threonine in the latter. The detection of 'pockets', often found to harbor active sites along the protein structure, identified 13 likely amino acid active sites in the GrpE protein for both *B. sporothermodurans* and *B. oleronius*. On the other hand, *B. cereus* and *B. subtilis* exhibited 38 and 12 pockets, respectively. The high numbers of likely amino acid active sites in *B. cereus* may be correlated with roles in enzymatic reactions involving its numerous virulence genes, since GrpE is associated with virulence [47].

5. Conclusions

Comparative analysis of *B. sporothermodurans* with other *Bacillus* species showed the close relatedness of these *Bacillus* species. Proteins implicated in heat resistance showed high levels of similarity in barring GrpE, which has unique substitutions that could infer different phenotypic responses to *B. sporothermodurans* with respect to heat resistance. These proteins, however, do not offer compelling differences between the HRS and non-HRS strains of *B. sporothermodurans*, which could unambiguously explain their specific characteristics regarding thermal resilience, although the possible involvement of the SpoIIP and ClpB can be highlighted. Further laboratory studies involving mutant and wild type strains of *B. sporothermodurans* are required to confirm the phenotypic effects of these genetic variations.

Author Contributions: Conceptualization, R.O.-D., S.D.d.O. and E.M.B.; Data curation, R.O.-D. and M.A.; Investigation, R.O.-D.; Methodology, R.O.-D., C.A.S.F., S.D.d.O. and E.M.B.; Resources, R.O.-D., M.A., A.I., S.D.d.O. and E.M.B.; Writing—original draft, R.O.-D.; Writing—review & editing, R.O.-D., C.A.S.F., S.D.d.O. and E.M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the South African Department of Science and Technology (DST)—National Research Foundation/Higher Education South Africa—India Brazil South Africa (NRF/HESA-IBSA-92056), in the form of a scholarship disbursed through the Center of Excellence (COE) in Food Security of the University of Pretoria, South Africa. This work was also supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES)—[Finance Code 001].

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- 1. Patel, S. Drivers of bacterial genomes plasticity and roles they play in pathogen virulence, persistence and drug resistance. *Infect. Genet. Evol.* **2016**, *45*, 151–164. [CrossRef] [PubMed]
- Magno-Perez-Bryan, M.C.; Martınez-Garcia, P.M.; Hierrezuelo, J.; Rodriguez-Palenzuela, P.; Arrebola, E.; Ramos, C.; de Vicente, A.; Perez-Garcia, A.; Romero, D. Comparative genomics within the Bacillus Genus reveal the singularities of two robust Bacillus amyloliquefaciens biocontrol strains. *MPMI* 2015, 28, 1102–1116. [CrossRef] [PubMed]
- Tettelin, H.; Riley, D.; Cattuto, C.; Medini, D. Comparative genomics: The bacterial pan-genome. *Curr. Opin. Microbiol.* 2008, 11, 472–477. [CrossRef] [PubMed]
- 4. Loman, N.J.; Pallen, M.J. Twenty years of bacterial genome sequencing. *Nat. Rev. Microbiol.* **2015**, *13*, 787–794. [CrossRef]
- Land, M.; Hauser, L.; Jun, S.R.; Nookaew, I.; Leuze, M.R.; Ahn, T.H.; Karpinets, T.; Lund, O.; Kora, G.; Wassenaar, T.; et al. Insights from 20 years of bacterial genome sequencing. *Funct. Integr. Genom.* 2015, 15, 141–161. [CrossRef]
- 6. Zwick, M.E.; Joseph, S.J.; Didelot, X.; Chen, P.E.; Bishop-Lilly, K.A.; Stewart, A.C.; Willner, K.; Nolan, N.; Lentz, S.; Thomason, M.K.; et al. Genomic characterization of the Bacillus cereus sensu lato species: Backdrop to the evolution of Bacillus anthracis. *Genome Res.* **2012**, *22*, 1512–1524. [CrossRef]

- McNally, A.; Oren, Y.; Kelly, D.; Pascoe, B.; Dunn, S.; Sreecharan, T.; Vehkala, M.; Välimäki, N.; Prentice, M.B.; Ashour, A.; et al. Combined Analysis of Variation in Core, Accessory and Regulatory Genome Regions Provides a Super-Resolution View into the Evolution of Bacterial Populations. *PLoS Genet.* 2016, 12, 1–16. [CrossRef]
- Alcaraz, L.; Moreno-Hagelsieb, G.; Eguiarte, L.E.; Souza, V.; Herrera-Estrella, L.; Olmedo, G. Understanding the evolutionary relationships and major traits of Bacillus through comparative genomics. *BMC Genom.* 2010, *11*, 1–17. [CrossRef]
- 9. Sharma, A.; Satyanarayana, T. Comparative Genomics of Bacillus species and its Relevance in Industrial Microbiology. *Genom. Insights* **2013**, *6*, 25–36. [CrossRef]
- 10. Nicholson, W.L.; Munakata, N.; Horneck, G.; Melosh, H.J.; Setlow, P. Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. *Microbiol. Mol. Biol. Rev.* 2000, 64, 548–572. [CrossRef]
- 11. Schallmey, M.; Singh, A.; Ward, O.P. Developments in the use of Bacillus species for industrial production. *Can. J. Microbiol.* **2004**, *50*, 1–17. [CrossRef] [PubMed]
- Scheldeman, P.; Herman, L.; Goris, J.; De Vos, P.; Heyndrickx, M.; De Vos, P.; Heyndrickx, M. Polymerase chain reaction identification of Bacillus sporothermodurans from dairy sources. *J. Appl. Microbiol.* 2002, *92*, 983–991. [CrossRef] [PubMed]
- Herman, L.M.F.; Vaerewijck, M.J.M.; Moermans, R.J.B.; Waes, G.M.A.V.J.; Waes, G.M.A.V.J. Identification and detection of Bacillus sporothermodurans spores in 1, 10, and 100 milliliters of raw milk by PCR. *Appl. Environ. Microbiol.* **1997**, 63, 3139–3143. [CrossRef] [PubMed]
- Heyndrickx, M.; Coorevits, A.; Scheldeman, P.; Lebbe, L.; Schumann, P.; Rodfiguez-Diaz, M.; Forsyth, G.; Dinsdale, A.; Heyrman, J.; Logan, N.A.; et al. Emended descriptions of Bacillus sporothermodurans and Bacillus oleronius with the inclusion of dairy farm isolates of both species. *Int. J. Syst. Evol. Microbiol.* 2012, 62, 307–314. [CrossRef]
- 15. Owusu-Darko, R.; Allam, M.; de Oliveira, S.D.; Ferreira, C.A.S.; Grover, S.; Mtshali, S.; Ismail, A.; Mallappa, R.H.; Tabit, F.; Buys, E.M. Genome Sequences of Bacillus sporothermodurans Strains Isolated from Ultra-High-Temperature Milk. *Microbiol. Resour. Announc.* **2019**, *8*, e00145-19. [CrossRef] [PubMed]
- 16. Owusu-Darko, R.; Allam, M.; Mtshali, S.; Ismail, A.; Buys, E.M. Draft genome sequence of Bacillus oleronius DSM 9356 isolated from the termite Reticulitermes santonensis. *Genom. Data* **2017**, *12*, 76–78. [CrossRef]
- Aziz, R.K.; Bartels, D.; Best, A.A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M.; et al. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genom.* 2008, 9, 75. [CrossRef]
- Overbeek, R.; Olson, R.; Pusch, G.D.; Olsen, G.J.; Davis, J.J.; Disz, T.; Edwards, R.A.; Gerdes, S.; Parrello, B.; Shukla, M.; et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 2014, 42, D206–D214. [CrossRef]
- 19. Brettin, T.; Davis, J.J.; Disz, T.; Edwards, R.A.; Gerdes, S.; Olsen, G.J.; Olson, R.; Overbeek, R.; Parrello, B.; Pusch, G.D.; et al. RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **2015**, *5*, 1–6. [CrossRef]
- Wattam, A.R.; Abraham, D.; Dalay, O.; Disz, T.L.; Driscoll, T.; Gabbard, J.L.; Gillespie, J.J.; Gough, R.; Hix, D.; Kenyon, R.; et al. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res.* 2014, 42, 581–591. [CrossRef]
- 21. Wattam, A.R.; Davis, J.J.; Assaf, R.; Boisvert, S.; Brettin, T.; Bun, C.; Conrad, N.; Dietrich, E.M.; Disz, T.; Gabbard, J.L.; et al. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Res.* **2017**, *45*, D535–D542. [CrossRef]
- 22. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]
- 23. Stamatakis, A. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* **2014**, *30*, 33–34. [CrossRef] [PubMed]
- 24. Rambaut, A. FigTree v1.4.3. Available online: https://tree.bio.ed.ac.uk/software/figtree/ (accessed on 6 February 2020).
- 25. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [CrossRef] [PubMed]
- 26. Hall, T. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.

- Capra, J.A.; Singh, M. Predicting functionally important residues from sequence conservation. *Bioinformatics* 2007, 23, 1875–1882. [CrossRef] [PubMed]
- 28. Kaas, R.S.; Leekitcharoenphon, P.; Aarestrup, F.M.; Lund, O. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS ONE* **2014**, *9*, 1–8. [CrossRef]
- 29. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The rapid generation of mutation data matrices. *Comput. Appl. Biosci.* **1992**, *8*, 275–282.
- 30. Felsenstein, J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* **1985**, *39*, 783. [CrossRef]
- Kuhnigk, T.; Borst, E.-M.; Breunig, A.; König, H.; Collins, M.D.; Hutson, R.A.; Kämpfer, P. Bacillus oleronius sp.nov., a member of the hindgut flora of the termite Reticulitermes santonensis (Feytaud). *Can. J. Microbiol.* 1995, 41, 699–706. [CrossRef]
- 32. Flint, S.; Gonzaga, Z.J.; Good, J.; Palmer, J. Bacillus thermoamylovorans—A new threat to the dairy industry—A review. *Int. Dairy J.* 2016, 65, 38–43. [CrossRef]
- 33. Yang, G.; Zhou, X.; Zhou, S.; Yang, D.; Wang, Y.; Wang, D. Bacillus thermotolerans sp. nov., a thermophilic bacterium capable of reducing humus. *Int. J. Syst. Evol. Microbiol.* **2013**, *63*, 3672–3678. [CrossRef] [PubMed]
- Verma, A.; Pal, Y.; Khatri, I.; Ojha, A.K.; Gruber-Vodicka, H.; Schumann, P.; Dastager, S.; Subramanian, S.; Mayilraj, S.; Krishnamurthi, S. Examination into the taxonomic position of Bacillus thermotolerans Yang et al., 2013, proposal for its reclassification into a new genus and species Quasibacillus thermotolerans gen. nov., comb. nov. and reclassification of B. encimensis Dastager et al. *Syst. Appl. Microbiol.* 2017, 40, 411–422. [CrossRef]
- 35. Bobay, L.M.; Ochman, H. The evolution of bacterial genome architecture. *Front. Genet.* **2017**, *8*, 1–6. [CrossRef] [PubMed]
- 36. Scheldeman, P.; Herman, L.; Foster, S.; Heyndrickx, M. Bacillus sporothermodurans and other highly heat-resistant spore formers in milk. *J. Appl. Microbiol.* **2006**, *101*, 542–555. [CrossRef] [PubMed]
- 37. Mcinerney, J.O.; Mcnally, A.; Connell, M.J.O. Why prokaryotes have pangenomes. *Nat. Publ. Gr.* 2017, 2, 1–5. [CrossRef] [PubMed]
- Mcinerney, J.O.; Mcnally, A.; Connell, M.J.O. Reply to 'The population genetics of pangenomes'. *Nat. Microbiol.* 2017, 2, 41564. [CrossRef]
- Raskin, D.M.; Seshadri, R.; Pukatzki, S.U.; Mekalanos, J.J. Bacterial genomics and pathogen evolution. *Cell* 2006, 124, 703–714. [CrossRef]
- LaBreck, C.J.; May, S.; Viola, M.G.; Conti, J.; Camberg, J.L. The Protein Chaperone ClpX Targets Native and Non-native Aggregated Substrates for Remodeling, Disassembly, and Degradation with ClpP. *Front. Mol. Biosci.* 2017, 4, 1–13. [CrossRef]
- 41. Bukau, B.; Horwich, A.L. The Hsp70 and Hsp60 chaperone machines. Cell 1998, 92, 351–366. [CrossRef]
- 42. Wickner, S.; Maurizi, M.R.; Gottesman, S. Posttranslational Quality Control: Folding, Refolding, and Degrading Proteins. *Science* 1999, *286*, 1888–1893. [CrossRef] [PubMed]
- Stoecklin, G.; Bukau, B. Telling right from wrong in life—Cellular quality control. *Nat. Rev. Mol. Cell Biol.* 2013, 14, 613. [CrossRef] [PubMed]
- 44. Beeby, M.; Connor, B.D.O.; Ryttersgaard, C.; Boutz, D.R.; Perry, L.J.; Yeates, T.O. The Genomics of Disulfide Bonding and Protein Stabilization in Thermophiles. *PLoS Biol.* **2005**, *3*, 1549–1558. [CrossRef]
- Ladenstein, R.; Ren, B. Protein disulfides and protein disulfide oxidoreductases in hyperthermophiles. *FEBS J.* 2006, 273, 4170–4185. [CrossRef] [PubMed]
- Hecker, M.; Schumann, W.; Völker, U. Heat-shock and general stress response in Bacillus subtilis. Mol. Microbiol. 1996, 19, 417–428. [CrossRef] [PubMed]
- 47. Liberek, K.; Georgopoulos, C. Autoregulation of the *Escherichia coli* heat shock response by the DnaK and DnaJ heat shock proteins. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11019–11023. [CrossRef]
- 48. Schroder, H.; Langer, T.; Hartl, F.; Bukaul, B. DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *EMBO J.* **1993**, *12*, 4137–4144. [CrossRef]
- 49. Gamer, J.; Multhaup, G.; Tomoyasu, T.; Mccarty, J.S.; Rudiger, S.; Schonfeld, H.-J.; Schirra, C.; Bujard, H.; Bukau, B. A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the Escherichia coli heat shock transcription factor sigma32. *EMBO J.* **1996**, *15*, 607–617. [CrossRef]
- 50. Lee, S.; Sowa, M.E.; Watanabe, Y.H.; Sigler, P.B.; Chiu, W.; Yoshida, M.; Tsai, F.T.F. The structure of ClpB: A molecular chaperone that rescues proteins from an aggregated state. *Cell* **2003**, *115*, 229–240. [CrossRef]

- 51. Doyle, S.M.; Wickner, S. Hsp104 and ClpB: Protein disaggregating machines. *Trends Biochem. Sci.* 2009, 34, 40–48. [CrossRef]
- 52. Kirstein, J.; Molière, N.; Dougan, D.A.; Turgay, K. Adapting the machine: Adaptor proteins for Hsp100/Clp and AAA+ proteases. *Nat. Rev. Microbiol.* **2009**, *7*, 589–599. [CrossRef] [PubMed]
- 53. Gerth, U.; Kirstein, J.; Mostertz, J.; Waldminghaus, T.; Miethke, M.; Kock, H.; Hecker, M. Fine-Tuning in Regulation of Clp Protein Content in Bacillus subtilis. *J. Bacteriol.* **2004**, *186*, 179–191. [CrossRef]
- 54. Glover, J.R.; Lindquist, S. Hsp104, Hsp70, and Hsp40: A novel chaperone system that rescues previously aggregated proteins. *Cell* **1998**, *94*, 73–82. [CrossRef]
- Tomoyasu, T.; Mogk, A.; Langen, H.; Goloubinoff, P.; Bukau, B. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the Escherichia coli cytosol. *Mol. Microbiol.* 2001, 40, 397–413. [CrossRef]
- 56. Schulz, A.; Schumann, W. hrcA, the first gene of the Bacillus subtilis dnaK operon encodes a negative regulator of class I heat shock genes. These include: hrcA, the First Gene of the Bacillus subtilis dnaK Operon Encodes a Negative Regulator of Class I Heat Shock Genes. J. Bacteriol. 1996, 178, 1088–1093. [CrossRef] [PubMed]
- 57. Hitomi, M.; Nishimura, H.; Tsujimoto, Y.; Matsui, H.; Watanabe, K. Identification of a helix-turn-helix motif of Bacillus thermoglucosidasius HrcA essential for binding to the CIRCE element and thermostability of the HrcA-CIRCE complex, indicating a role as a thermosensor. *J. Bacteriol.* **2003**, *185*, 381–385. [CrossRef]
- 58. Chastanet, A.; Losick, R. Engulfment during sporulation in Bacillus subtilis is governed by a multi-protein complex containing tandemly acting autolysins. *Mol. Microbiol.* **2007**, *64*, 139–152. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).