Journal of Global Biosciences
ISSN 2320-1355
Volume 6, Number 6, 2017, pp. 5060-5069
Website: www.mutagens.co.in

Research Paper

PATTERN OF HETEROGENEITY IN A CELL WALL-ASSOCIATED PROTEIN A (wprA) AMONG THE *Bacillus* STRAINS

Ikram-Ul Haq, Zanobia, Nazia Parveen Gill and Uzera Syed[1]

Institute of Biotechnology and Genetic Engineering (IBGE),
University of Sindh, Jamshoro-76080, Pakistan,
[1]Department of Statistics,
University of Sindh, Jamshoro-76080,,
Pakistan.

Abstract

In this taxonomic study, a phylogenetic tree was constructed and disparity index also computed from the online available wprA (wall-associated protein A; EC 3.4.21) polypeptide sequences in some *Bacillus* strains. The wprA protein is being conserved region of core genome of *Bacillus* genus, while its presence or absence among other non-*Bacillus* bacterial strains is result of horizontal gene transfer. From phylogenetic analysis, it is observed that sequence homology of wprA is variable among *Bacillus* strains. Its sequences homology is not stable because of bacterial mode of nutrition and habituation or developed through transcriptional errors due to breakage of single DNA strand when collision of polymerase occurs. Such specific pattern of wprA distribution in bacterial population may be a best approach for comparative maximum likelihood analysis for the concatenated alignment of other conserved single allele genes in bacterial population.

Key words: *Bacillus subtilis*; *proteases*; wprA; molecular evolution, phylogenetic tree, Disparity Index.

INTRODUCTION

Genomics as well as proteomics database of more than 3000 microbial strains of bacterial and archaeal taxa are available to examine their evolutionary trend from time to time as well as habitat to habitat. Utility of their polypeptide sequence based data is valuable for the understanding of how are they inter-related phylogenetically. Recently, it is assumed that evolution has been strictly bifurcating tree like processes among the living organisms [1,2]. However, such arguments, a tree-like depiction especially for microbes is not valid due to horizontal gene transfer from one lineage to another. It can obscure the chance of vertical line of descent among species [3,4]. Horizontal gene transfer remains important in understanding evolutionary biology of a specie while some genes shows lowest rate than others [5]. Definitely, such potential horizontal gene transfer are useful in the construction of a framework of vertical inheritance lineage

among the microbial species [6,7]. Although the phylogenetic trees are useful tools for the guidance to select a particular genomes for sequencing [8], to assign a taxonomic community sequence data [9], identification of ecological adaptations [10] inferring of biogeographical co-speciation [11], to perform the analysis of phylogenetic profiles of the arising species [12] and may be more.

Most species of genus *Bacillus* and Firmicutes are gram positive, endosporic, rod-shaped non-pathogenic bacteria commonly found in diverse places like as clay soil, rocks, dust, aquatic places, in plant vegetation, in food and even in gastrointestinal tracts of living organisms [13]. Its cell wall is comprised on higher proportion of lipoproteins. Relatively, cell wall of Gram positive bacteria is thick due to peptidoglycan rich and complex structure linked covalently with other anionic polymers [14]. For being its electrostatic nature, a large number of active proteins are attached in cell wall of *Bacillus* specie. With proteinaceous paracrystalline S-layers on the outer surface of wall has remarkable resistance against high pressure, temperature or rapid surface tension [15,16]. Such nature of cell wall is most likely to require for activation of peptidoglycan degrading enzymes that are essential even for cell growth [17-20]. Among other peptidoglycan hydrolases, the wprA (cell wall-associated protein A) is also one of the most important protease. In actual, main three extracytoplasmic serine proteases (wprA, htrA and htrB) acts as 'quality control' of secretory proteins through cell membrane to wall interface, while degrade the misfolded ones. These proteases are activated through signal transduction pathways in response to secretion of misfolded proteins or other physiological (e.g. heat) stresses. Even the presence of these proteases remains the major barrier to many heterologous proteins secretions at pharmaceutical level by gram-positive bacteria [21,22].

The wprA (~96 kDa) belongs to monocistronic operon and express during exponential growth phase [21,22]. The *Bacillus* cell wall has been predicted as being large extracytoplasmic serine proteases. No any growth phenotype was affected with wprA-null mutation while production of native-wall proteins and some heterologous proteins secretion was enhanced [22]. The WprA also involved in thermal sensitivity inactivation which is regulated by two component (YvrG-YvrHb) signal transduction pathways of major cell-wall autolysins [23]. Meanwhile, induction of these signals not clear either wprA is activated in response to protein secretion stress.

For ancestor analysis of a specific protein in bacterial population various methods has been applied while each method is depended on phylogenic relationships to test the rising evolutionary hypotheses. For the lineage analysis, the phylogenetic analysis tools are focused on extant organisms as well as the information from extinct taxa or fossils can be incorporated. In general comparative phylogenetic approaches can be grouped into two types like as phenotypic or genetic (dependent on some internal individual potential of some characters that infer the evolutionary history across a phylogeny) and diversification rates (external affects faced by individual that infer in evolutionary branching within a phylogeny. Even some phylogenetic approaches that can do both simultaneously. In typical to use protein data set to develop a tree in conjunction to find the comparative phylogenetic relationships among lineages with their time era (length of branches) can be computed easily.

In this study, a phylogenic relationship among *Bacilli* and other bacterial species was sketched on the basis of single cell wall bounded protein known as wprA. Almost all bacterial wprA polypeptides have showed heterozygous form. The aim of this study is to simply develop a phylogenetic tree of *Bacillus* strains on the basis of amino acid sequence of a single protein. The possible evolution among collected bacterial strains

was analyzed with two different softwares or Databases (BLAST (from NCBI) and MEGA version 7.0). Although functionally as well as genetically wprA appears to share their same evolutionary origin, while differ in protein size and functional liabilities under certain circumstances.

MATERIALS AND METHODS

The cell cultures of selected bacterial strains were grown from a single colony. These colonies were raised from the glycerol stocks on LB (Luria-Bertani) agar plates. The culture was incubated for over-night (O.N.) on rotating incubator (37°C, 250 rpm). The 1.0 ml O.N. culture was transferred to 100 ml fresh LB broth and incubated for 1 hour again on rotating incubator until cell culture OD_{600} reaches to 0.5. Cells were collected with centrifuge the LB broth in 1.5 ml eppendorf tubes at 10,000 rpm for 5 minutes. Cell pallet was washed with dH_2O for 3-times and mixed in 50 µl cell lysis buffer (0.06 M Tris-HCl, 2.5% Glycerol, 0.5% SDS, 1.25% ß- mercaptoethanol). These samples were denatured by heating at 95°C for 5-10 minutes. The solubilized protein samples were fractionated with 12.5% SDS-PAGE resolving gel at 30 mA for 40 min [24]. The Gel was stained with Coomassie Brillant Blue (R-250) and its photo was captured.

A single polypeptide namely wprA (cell wall-associated protein A) and its amino acid sequences among 10 closely but heterozygous *Bacillus* strains were selected. These polypeptides with their accession #s were tabulated from online available NCBI protein database (http://www.ncbi.nlm.nih.gov /protein/) and Uni-protein data-set (http://www. uniprot.org/). These selected *Bacillus* strains were tabulated as shown in table 1 in correspond to their accession numbers, size of polypeptide and their level of homozygosity.

For the wprA maximum likelihood UPGMA-tree, the targeted wprA sequences of *Bacillus* strains were submitted for comparative inference of evolutionary mode with UPGMA method [25]. The UniProtein data-set (Table 1) of *Bacillus* species was submitted for maximum likelihood relationship with JTT (Jones-Taylor-Thornton) or WAG (Whelan and Goldman) matrix-based model [26,27]. For optimal tree sum of branch length was 3.16977603 with 50 replicates [27]. The *p*-distance method was employed (amino acid differences per site) to compute evolutionary distance [28]. Total 888 positions were found in total dataset and its analysis was based on 10 amino acid sequences (with elimination of gaps and missing data). The above evolutionary analyses were performed with MEGA7 software [29].

The wprA sequences were also submitted for the estimation of homogeneity of substitution pattern. The probability of evolution in polypeptide sequence pattern of substitution (Disparity Index test) were figured out and their *p*-values were estimated by using Monte-Carlo method (Table 2, yellow high-lighted diagonals) [30]. The estimation disparity index per site involved on 11 amino acid sequences with elimination of gaps and missing data. Among the polypeptide sequences total of 885 positions were targeted in final dataset [29]. It was also calculated by using MEGA (7.0) software.

RESULTS AND DISCUSSION

In systematic biology, comparative studies among the variant observable features or phenotypes has been intensely studied in inter or intra-species [31] either in wild and mutant lines [32] and even human genetic phenotypes [33]. For being fundamental importance, the procedures are experimental research based of genetics for the developing living organism or its growth relations with other individuals. Such available

phenotypic as well as genotypic data remain useful for the computational based determination of evolutionary processes or event going on in nature. The experimental procedures, although expresses a biological form or a phenotype precisely, while serious limitations in consensus to their comparative integrating data across the phenotype to genetic databases remain hidden.

The *Bacillus* and other related bacterial species have been widely used for the production of a large number of heterologous proteins including many industrially important enzymes useful in the processing of leather, detergent, food and beverages. These appropriate proteins are not only produced by them but they are also capable to secrete these enzymes at higher concentrations in its cell culture medium. Purification of secreted proteins from the cell culture is easiest and cheapest than from cytoplasm of bacteria. These organisms are also reducing the down-stream processing costs of the produced proteins. From the last few decades, a number of researchers cited their considerable efforts that have been aimed to develop *B. subtilis* as a horse-host for bio-production of heterologous proteins. The systematic ways from protein synthesis to culture medium exportation, folding of exported protein have a key importance for its future fate. The correct folding of polypeptide is dependent on the cell wall bounded proteases [22,34]. The extra-cellular including wall associated *proteases* are very important for the elimination of abnormal or non-structured proteins. The wprA is a quality controller of secreting heterologous proteins. It is being single polypeptide and found variant or could be of multiple forms among the *Bacillus* strains.

In according to available online database, 11 distinct *Bacillus* species were selected on the basis of presence wprA *protease*. Phylogenetically they could be divided into 11 genetically distinct lineages. These polypeptide sequences were used to build a phylogenetic tree (Table 1, Fig 2) for evolutionary studies. This analysis could enable us to judge more precise evolutionary and phylogenetic relationships among these *Bacillus* lineages. Amino acids composition between polypeptides sequences, over all mean, divergence or distance from one specie to specie has been observed.

The wprA polypeptide was observed on SDS-PAGE (Fig 1). It showed almost similar size among the *Bacillus* species from 95kDa to 113kDa. The longest polypeptide with size 148kDa was synthesized by *Salmonella enterica*. The indicated bands marked in figure 1 are exactly of similar sized in corresponds to the online available database of wprA of *Bacillus* strains. Meanwhile, the largest wprA in *Salmonella* species could be of great valuable in commercial point of view for being as a quality controller for heterologous protein secretion rates from the cell to medium. Variations among the lengths of wprA polypeptides from bacterial species to species may lead to have a difference in their activities or barrier in exporting proteins through cell wall. It could be important for being stable against extreme environmental conditions. Such variation in physical and physiological as well as in chemical structures of the single polypeptide among bacterial species could be subject for systematic biology for the study evolutionary processes.

In according to the data calculated through disparity index of substitution pattern from the wprA polypeptide sequences of the selected bacterial strains each showed certain level of heterogenity (Table 2). The values assumed with Disparity Index (DI) are observed equal to zero means that the bacterial species are endangered and demanding for their improvement. If DI is more than zero while lesser than 1 means that species are under evolutionary processes and getting the level of stability at speciefic habitate. Similarly, DI is higher than 1 means that species is exceeding to adopt itself with the changing environment. The DI results are confirming the presence of variation in the polypeptide sequences, which is also being a good indiction for the creation of new

species or even their distiniction. The substitution pattern indicate the phylogenetic evolution among the selected strains (Fig 2). This variant adoptation could be need of the habitate wherethere these strains are surviving or modyfying itself for their survival. The wprA polypeptide size specification of each strain indicates the species divergence or creation of a species from time to time. This divergence is not dependent on addition or deletion of certain genome sequences linked to wprA. It is occurring at the whole genome level. Similarly, the observed sequence divergences are not being a reflection of high rate of point mutation when the case is heterogeneous lineages [35,36].

CONCLUSIONS

In molecular taxonomy, phylogenetic studies of DNA as well as polypeptide sequences have played very important role for the analysis of evolution from lower to higher organisms. Recent studies suggest that the protein based trees have been considered as being more reliable in some cases of phylogenetic trees based analyses. On the basis of this opinion, wprA polypeptide among different *Bacillus* strains used for the study of Disparity Index and phylogenetic based tree studies. The wprA is present among the bacterial strains in multiple allelic forms. On the basis of polypeptide sequences, the MEGA 7.0 software bi-furcates these 10 *Bacillus* strains into two major phylogenetic groups. Each variant species could be happened because of continuous alterations in genome due to substitutions or deletions of DNA sequences or even in addition to the occurrences of multiple point mutations. The efficacy of wprA could be variable from species to species as it is being most important cell wall associated protease that is available as quality controller of extra-cellular protein when exporting through the cell wall from cytoplasm to medium. For future work, it could be helpful for commercial point of view to determine the suitability of wprA in favor to secretions of important pharmaceutical proteins.

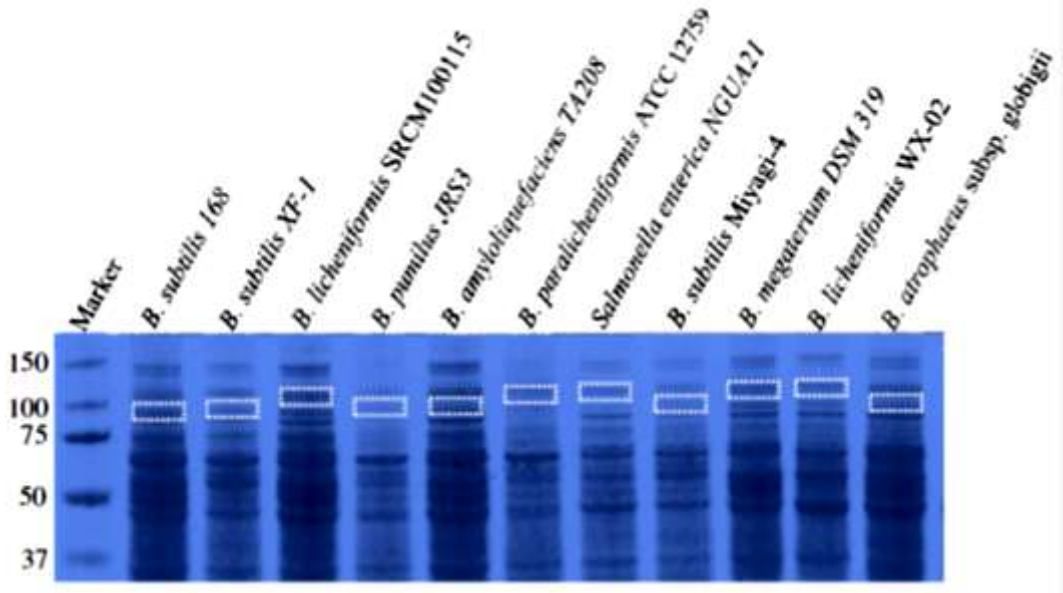

Figure 1. The SDS-PAGE with fractionated proteins of cell pallet of different *Bacillus* strains. The bands of wprA were marked in according to its relatively possible location on the gel. Lane 1: Protein marker (BioRad); **Lane 02 to 12**: Fractionated the cell pallet proteins with expected wprA protein of different bacterial culture like as *Bacillus subtilis* 168 (96.488 kDa); *Bacillus subtilis* XF-1 (95.439 kDa);

Bacillus licheniformis (112.130 kDa); *Bacillus pumilus* (96.484 kDa); *Bacillus amyloliquefaciens* (95.906 kDa); *Bacillus paralicheniformis* (112.048 kDa); *Salmonella enterica* (148.029 kDa); *Bacillus subtilis* Miyagi-4 (96.499 kDa); *Bacillus megaterium* DSM 319 (113.857 kDa); *Bacillus licheniformis* WX-02 (112.147 kDa); *Bacillus atrophaeus* subsp. globigii (95.337 kDa).

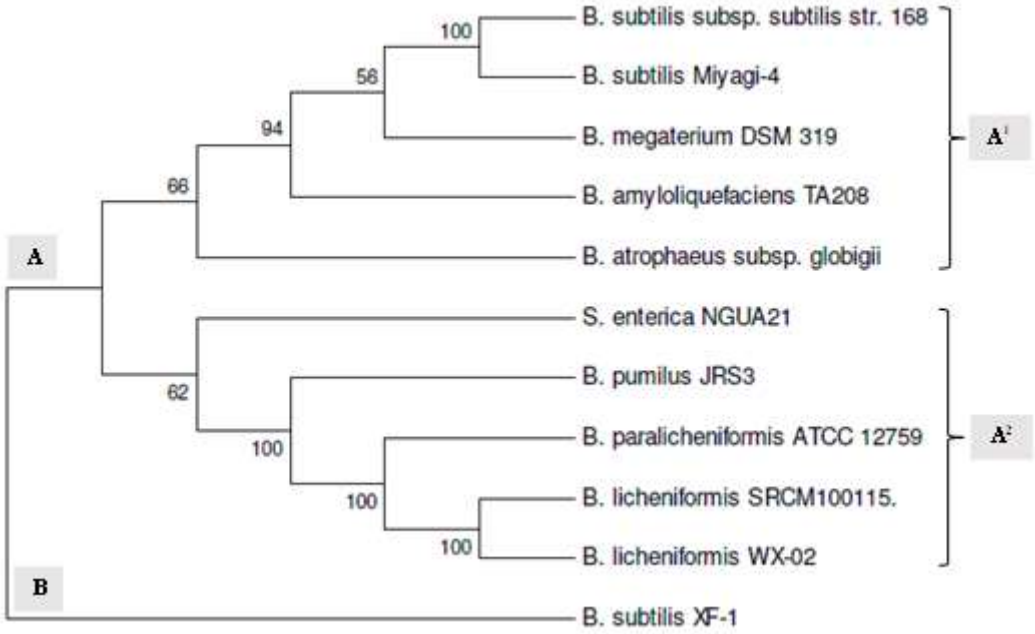

Figure 2. Evolutionary relationships of among *Bacillus* strains in maximum likelihood tree of wprA polypeptide sequences. In tree explorer window the round circle marked wprA polypeptide or gene duplications and wprA's speciation events marked with triangle shapes.

Table 1. Comparative wprA polypeptide and its amino acid sequence homology in *Bacillus* strains

#s	Accession #	Bacterial strains	Polypeptides	Max Scores	Total Scores	Query cover %	Identical %	References
a.	CAB12917.1	*B. subtilis* subsp. subtilis str. 168	894 aa	1803	1803	100	100	[36]
b.	AGE62929	*Bacillus subtilis* XF-1	885 aa	1771	1771	98	99	[37]
c.	OAZ65519.1	*B. licheniformis* SRCM100115	1047 aa	800	969	99	50	[38]
d.	CUB15022.1	*B. pumilus* JRS3	888 aa	848	848	98	50	[39]
e.	AEB23192.1	*B. amyloliquefaciens* TA208	891 aa	1274	1274	100	71	[40]
f.	KFM90595.1	*B. paralicheniformis* ATCC 12759	1047 aa	812	983	99	50	[41]
g.	GAR87140.1	*S. enterica* NGUA21	1264 aa	49.7	150	16	28	[42]
h.	GAK79780.1	*B. subtilis* Miyagi-4	894 aa	1784	1784	100	99	[43]
i.	ADF38014.1	*B. megaterium* DSM 319	1066 aa	848	1019	100	50	[44]
j.	AKQ74115	*B. licheniformis* WX-02	1047 aa	801	971	99	50	[45]
k.	AIK46855	*B. atrophaeus* subsp. globigii	893 aa	1063	1063	99	60	[41]

Table 2. Disparity index of substitution patterns among the polypeptide sequences of wprA in *Bacillus* strains.

Bacterial strains	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
01. *B. subtilis* subsp. subtilis str. 168		0.000	0.017	0.000	0.000	0.000	8.119	0.000	0.118	0.000	0.000
02. *Bacillus subtilis* XF-1	1.000		0.182	0.000	0.000	0.000	8.252	0.000	0.249	0.000	0.000
03. *B. licheniformis* SRCM100115	0.400	0.000		0.000	0.000	0.000	5.380	0.031	0.000	0.000	0.000
04. *B. pumilus* JRS3	1.000	0.000	0.000		0.000	0.000	8.069	0.000	0.131	0.000	0.000
05. *B. amyloliquefaciens* TA208	1.000	0.000	0.000	0.000		0.000	7.238	0.000	0.000	0.000	0.000
06. *B. paralicheniformis* ATCC 12759	1.000	0.000	0.000	0.000	0.000		5.671	0.000	0.000	0.000	0.000
07. *S. enterica* NGUA21	0.000	0.000	0.000	0.000	0.000	0.000		8.195	5.626	5.246	5.460
08. *B. subtilis* Miyagi-4	1.000	1.000	0.420	1.000	1.000	1.000	0.000		0.166	0.097	0.166
09. *B. megaterium* DSM 319	0.314	0.182	1.000	0.264	1.000	1.000	0.000	0.258		0.000	0.000
10. *B. licheniformis* WX-02	0.334	0.238	1.000	1.000	1.000	1.000	0.000	0.300	1.000		0.000
11. *B. atrophaeus* subsp. globigii	0.348	0.228	1.000	1.000	1.000	1.000	0.000	0.286	1.000	1.000	

The amino acids divergences per site out of 885 positions are based on a site with 11 amino acid sequences. The homogeneity was analyzed with MEGA 7.0 [29] *to estimate p-values* [30].

REFERENCES

[1] Felsenstein, J. **2004**, Inferring Phylogenies. Am. J. Hum. Genet., 74(5) pp.1074.
[2] Fitch, W.M., and Margoliash, E., **1967**, Construction of phylogenetic trees. Science, 155 pp.279-284.
[3] Hilario G.J., **1993**, Horizontal transfer of ATPase genes the tree of life becomes a net of life. Biosystems, 31, pp.111-119.
[4] Dagan, T., Artzy-Randrup, Y., and Martin, W., **2008**, Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. Proc. Natl. Acad. Sci. U. S. A., 105, pp.10039-10044.
[5] Hao, W., and Golding, G.B., **2008**, Uncovering rate variation of lateral gene transfer during bacterial genome evolution. BMC Genomics, 9, pp.235.
[6] Puigbò, P., Wolf, Y.I., and Koonin, E.V., **2009**, Search for Tree of Life in the thicket of the phylogenetic forest. J Biol, 8, pp. 59.
[7] Leigh, J.W., Schliep, K., Lopez, P., and Bapteste, E., **2011**, Let them fall where they may: Congruence analysis in massive phylogenetically messy data sets. Mol. Biol. Evol., 28, pp.2773-2785.
[8] Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N.N., Kunin, V., Goodwin, L., Wu, M., Tindall, B.J., Hooper , S.D., Pati, A., Lykidis, A., Spring, S., Anderson, I.J., D'haeseleer. P., Zemla, A., Singer, M., Lapidus, A., Nolan, M., Copeland, A., Han, C., Chen, F., Cheng, J.F., Lucas, S., Kerfeld, C., Lang, E., Gronow, S., Chain, P., Bruce, D., Rubin, E.M., Kyrpides, N.C., Klenk, H.P., and Eisen, J.A., **2009**, A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. Nature, 462, pp.1056-1060.
[9] Wu, M.and Eisen, J.A., **2008**, A simple, fast, and accurate method of phylogenomic inference. Genome Biol., 9, pp.151.
[10] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J., **1990**, Basic Local Alignment Search Tool. J. Mol. Biol., 215, pp.403-410.
[11] Sandström, J.P., Russell, J.A., White, J.P., and Moran, N.A., **2001**, Independent origins and horizontal transfer of bacterial symbionts of aphids. Mol. Ecol., 10, pp.217-228.
[12] Wu, M. et al., **2005**, Life in hot carbon monoxide: The complete genome sequence of carboxydothermus hydrogenoformans Z-2901. PLoS Genet., 1, pp.563-574.
13] Nicholson, W.L., **2002**, Roles of Bacillus endospores in the environment. Cell. Mol. Life Sci., 59, pp.410-416.
[14] Perkins, H.R., Rogers, H.J., and Ward, J.B., **1980**, Microbial Cell Walls and Membranes. Chapman & Hall, London.
[15] Kembel, S.W., Eisen, J.A., Pollard, K.S., and Green, J.L., **2011**, The phylogenetic diversity of metagenomes. PLoS One, 6, pp.8.
[16] Kuroda, A., and Sekiguchi, J., **1990**, Cloning, sequencing and genetic mapping of a *Bacillus subtilis* cell wall hydrolase gene. J. Gen. Microbiol., 136, pp.2209-2216.
[17] Herbold, D.R., and Glaser, L., **1975**, Interaction of N-acetylmuramic acid L-alanine amidase with cell wall polymers. J. Biol. Chem., 250, pp.7231-7238.
[18] Regamey, A., and Karamata, D., **1998**, The n-acetylmuramoyl-l-alanine amidase encoded by the Bacillus subtilis 168 prophage SP. Microbiology, 144, pp.885-893.
[19] Smith, T.J. and Foster, S.J., **1995**, Characterization of the involvement of two compensatory autolysins in mother cell lysis during sporulation of Bacillus subtilis 168. J. Bacteriol., 177, pp.3855-3862.
[20] Darmon, E., et al., **2002**, A novel class of heat and secretion stress-responsive

genes is controlled by the autoregulated CssRS two-component system of *Bacillus subtilis*. J. Bacteriol., 184, pp.5661-5671.

[21] Harwood, C.R., and Cranenburgh, R., **2008**, *Bacillus* protein secretion: an unfolding story. Trends Microbiol., 16, pp.73-79.

[22] Stephenson, K., and Harwood, C.R., **1998**, Influence of a cell-wall-associated protease on production of alpha-amylase by *Bacillus subtilis*. Appl. Environ. Microbiol., 64, pp.2875-2881.

[23] Serizawa, M., Kodama, K., Yamamoto, H., Kobayashi, K., Ogasawara, N., and Sekiguchi, J., **2005**, Functional analysis of the YvrGHb two-component system of Bacillus subtilis: identification of the regulated genes by DNA microarray and northern blot analyses. Biosci. Biotechnol. Biochem., 69, pp.2155-2169.

[24] Laemmli, U.K., **1970**, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227, pp.680-685.

[25] Sneath, P.H.A., and Sokal, R., **1962**. Numerical Taxonomy. Nature, 193, pp.855-860.

[26] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S., **2011**, MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony. Mol. Biol. Evol., doi: 10.1093/molbev/msr121.

[27] Felsenstein, J., **1985**, Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 39, pp.783-791.

[28] Nei, M., and Kumar, S., **2000**, Molecular Evolutionand Phylogenetics, pp.*154*-158. New York: Oxford University Press.

[29] Kumar, S., Stecher, G., and Tamura, K., **2016**, MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol., pp.1870-1874.

[30] Kumar, S., and Gadagkar, S.R., **2001**, Disparity index: A simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. Genetics, 158, pp.1321-1327.

[31] Brochu, C.A., and Sumrall, C.D., **2001**, Phylogenetic Nomenclature and Paleontology. J. Paleontol., 75, pp.754-757.

[32] Blake, J.A., Bult, C.J., Kadin, J.A., Richardson, J.E., and Eppig, J.T., **2011**, The mouse genome database (MGD): Premier model organism resource for mammalian genomics and genetics. Nucleic Acids Res., 39, pp.D842-D848.

[33] Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., and McKusick, V.A., **2005**, Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucl. Acid. Res., 33, pp.D514-D517.

[34] Tamura, K., and Kumar, S., **2002**, Evolutionary distance estimation under heterogeneous substitution pattern among lineages. Mol. Biol. Evol., 19, pp.1727-1736.

[35] Sheffield, N.C., **2013**. The Interaction between Base Compositional Heterogeneity and Among-Site Rate Variation in Models of Molecular Evolution. Evol. Biol., 2013, pp.1-8.

[36] Kunst, F., et al., **1997**, The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature, 390, pp.249-256.

[37] Guo, S., Li, X., He, P., Ho, H., Wu, Y., and He, Y., **2015**, Whole-genome sequencing of Bacillus subtilis XF-1 reveals mechanisms for biological control and multiple beneficial properties in plants. J. Ind. Microbiol. Biotechnol., 42, pp.925-937.

[38] Song, Y.R., **2016**, Genome sequencing of Bacillus licheniformis strain SRCM100115. *Unpublished*.

[39] Abo-Aba, S.E.M., et al., **2015**, Draft Genome Sequence of Bacillus Species from the Rhizosphere of the Desert Plant Rhazya stricta. Genome Announc., 3, pp.e00957-15.

[40] Zhang, G., Deng, A., Xu, Q., Liang, Y., Chen, N., and Wen, T., **2011**, Complete genome sequence of Bacillus amyloliquefaciens TA208, a strain for industrial production of guanosine and ribavirin. J. Bacteriol., 193, pp.3142-3143.

[41] Ash, C., Jae, F., Wallbanks, S., and Phylogenetic, C., **2014**, Twenty Whole-Genome. 2, pp.4-5.

[42] Useh, N.M., et al., **2016**, Draft Genome Sequences of 37 Salmonella enterica Strains Isolated from Poultry Sources in Nigeria. Genome Announc., 4, pp.e00315-16.

[43] Hachiya, Y.T., Miyake, M., Hase, S., Kubo, Y., Kimura, K., and Sakakibara, **2017**, Draft genome sequence of Bacillus subtilis Miyagi-4, an efficient starter strain for natto (fermented soybean) production. *Unpublished*.

[44] Eppinger, M., et al., **2011**, Genome sequences of the biotechnologically important Bacillus megaterium Strains QM B1551 and DSM319. J. Bacteriol., 193, pp.4199-4213.

[45] Yangtse, W., et al., **2012**, Genome sequence of Bacillus licheniformis WX-02. J. Bacteriol., 194, pp.3561-3562.