Characterization and Comparative Genomics Analysis of two *Bacillus megaterium* lytic bacteriophages

Next Generation Sequencing (NGS) technologies provide unique possibilities for the comprehensive assessment of the environmental diversity of bacteriophages. Several *Bacillus* bacteriophages have been isolated, but very few *Bacillus megaterium* bacteriophages have been characterized. In this study, we describe the biological characteristics, whole genome sequences, and their annotations for two new isolates of the *B. megaterium* bacteriophages (BM5 and BM10), which were isolated from Egyptian soil samples. Growth analyses indicated that the phages BM5 and BM10 have a shorter latent period (25 and 30 minutes respectively) and a smaller burst size (103 and 117 PFU respectively), in comparison to that which is typical for *Bacillus* phages. The genome sizes of the phages BM5 and BM10 were 165,031 bp and 165,213 bp, respectively, with a modular organization. Bioinformatic analyses of these genomes enabled the assignment of putative functions to 97 and 65 putative ORFs, respectively. Comparative analysis of the BM5 and BM10 genome structures, in conjunction with other *B. megaterium* bacteriophages, revealed relatively high levels of sequence and organizational identity. Both genomic comparisons and phylogenetic analyses support the conclusion that the sequenced phages (BM5 and BM10) belong to different sub-clusters (L5 and L7 respectively), within the L-cluster, and display different lifestyles (lysogenic and lytic respectively). Moreover, sequenced phages encode proteins associated with *Bacillus* pathogenesis. In addition, BM5 does not contain any tRNA sequences, whereas BM10 genome codes for 17 tRNAs.

Characterization and Comparative Genomics Analysis of two *Bacillus megaterium* lytic Bacteriophages

Sharaf A.[1,2], Oborník M.[2,3], Hammad A.[4], El-Afifi S.[5], Marei E.[5]

[1] Genetic Depart., Faculty of Agriculture, Ain Shams Uni., Egypt. [2] Biology Centre CAS, Institute of Parasitology, Czech Republic. [3] Faculty of Science, University of South Bohemia, Czech Republic. [4] Microbiology Depart., Faculty of Agriculture, Minia Uni., Egypt. [5] Microbiology Depart., Faculty of Agriculture, Ain Shams Uni., Egypt.

Abstract

Several *Bacillus* bacteriophages have been isolated, but very few *Bacillus megaterium* bacteriophages have been characterized. In this study, we describe the biological characteristics, whole genome sequences, and their annotations for two new isolates of the *B. megaterium* bacteriophages (BM5 and BM10), which were isolated from Egyptian soil samples. Growth analyses indicated that the phages BM5 and BM10 have a shorter latent period (25 and 30 minutes respectively) and a smaller burst size (103 and 117 PFU respectively), in comparison to that which is typical for *Bacillus* phages. The genome sizes of the phages BM5 and BM10 were 165,031 bp and 165,213 bp, respectively. Comparative analysis of the BM5 and BM10 genome structures, in conjunction with other *B. megaterium* bacteriophages, revealed relatively high levels of sequence and organizational identity. Both genomic comparisons and phylogenetic analyses support the conclusion that the sequenced phages (BM5 and BM10) belong to different sub-clusters (L5 and L7 respectively), within the L-cluster, and display different lifestyles (lysogenic and lytic respectively). Moreover, sequenced phages encode proteins associated with *Bacillus* pathogenesis. In addition, BM5 does not contain any tRNA sequences, whereas BM10 genome codes for 17 tRNAs.

Conclusion

The present study reports the biological and genome properties of virulent and temperate *B. megaterium* phages. Both the BM5 and BM10 phages lacked repressor determinants to maintain lysogeny by down-regulating lytic promoters and to confer superinfection immunity. This increases its potential risk with respect to the use of *B. megaterium* as a biocontrol and a biofertilizer agent. We also present the first comparative whole genome nucleotide sequences analysis and a large terminase (TerL) protein phylogeny, thereby revealing clustering and sub-clustering of *B. megaterium* phages. Moreover, putative tRNAs were identified, revealing the ability of the BM10 phage to infect other hosts.

Furthermore, this study presents the most convincing phylogenetic analysis of the *Bacillus* phages, based on terminase amino acids sequences, exhibiting a robust relationship between the phage families and packing strategies, thus supporting the fact that the distribution of tRNA genes in the *Bacillus* phage is sub-cluster specific. On the other hand, our screening of the 87 *Bacillus* genomes phages for integrases and tRNA genes revealed that 66.7% of the 66 lytic *Bacillus* phages lack tRNA genes, whereas 25% of the 20 temperate *Bacillus* phages contained tRNA genes. Hence, we do not agree with the assumption that lytic bacteriophages are more likely to contain tRNA genes in comparison to temperate bacteriophages.

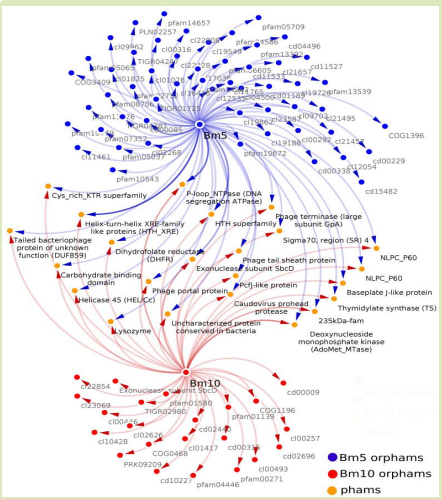

Fig.: Related Conserved Domains (CD) to the detected Phamilies (phams) using Phamerator. Both BM5 and BM10 orphams (phams containing a gene product from a single phage) presented as blue and red nodes, respectively. While shared phams (orange nodes) located between them. All nodes connected by blue and red arrows represent genes encode for each pham.

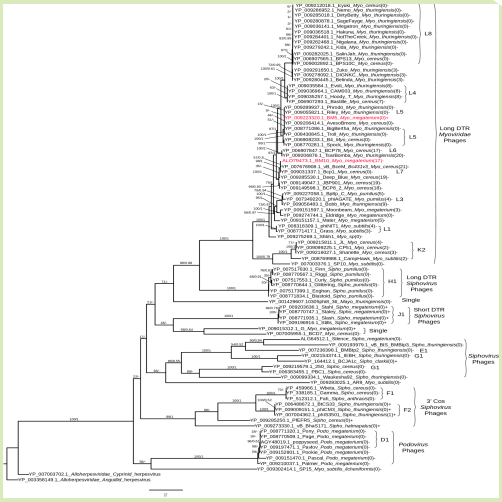

Fig.: Phylogenetic tree analysis of the *Bacillus* terminase reflects complete genome cluster. Sequenced phages colored in red, whole genome sub-cluster assignment and the phage families indicated on the right. Replication approaches for phages also indicated as, a direct terminal repeats (DTR) and cohesive ends (cos). Moreover, the presence and absence of integrase genes as (+) or (-), respectively and number of tRNA genes for each phage were indicated as the number between brackets. The phylogenetic tree constructed using a Maximum Likelihood distance matrix **(Stamatakis, 2014)** and Bayesian Monte Carlo Markov Chain (MCMC) matrix methods **(Lartillot, Lepage & Blanquart, 2009)**. the bootstraps from both methods merged in one rooted tree. Bootstrap values indicate the number of times a node was supported in 2000 re-sampling replications.

Background

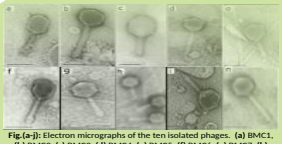

Fig.(a–j): Electron micrographs of the ten isolated phages. **(a)** BMC1, **(b)** BMC2, **(c)** BMC3, **(d)** BMC4, **(e)** BMC5, **(f)** BMC6, **(g)** BMC7, **(h)** BMC8, (i) BMC9 and (j) BMC10. **(Elmaghraby *et al.*, 2015)**

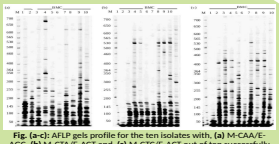

Fig. (a–c): AFLP gels profile for the ten isolates with. **(a)** M-CAA/E-AGC, **(b)** M-CTA/E-ACT and **(c)** M-CTC/E-ACT out of ten successfully primer combinations. **(Elmaghraby *et al.*, 2015)**

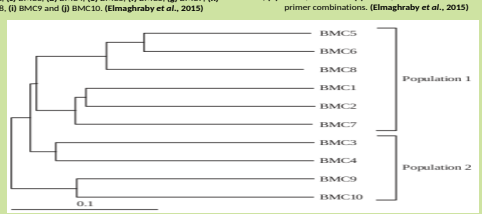

Fig.: UPGMA tree based on 524 AFLP loci used to genotype 10 isolates of phages specific for *Bacillus megaterium* belonging to two different populations (host specificity) **(Elmaghraby *et al.*, 2015)**

Results

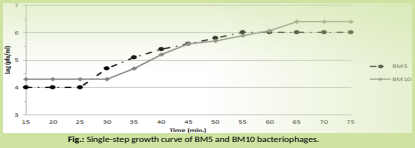

Fig.: Single-step growth curve of BM5 and BM10 bacteriophages.

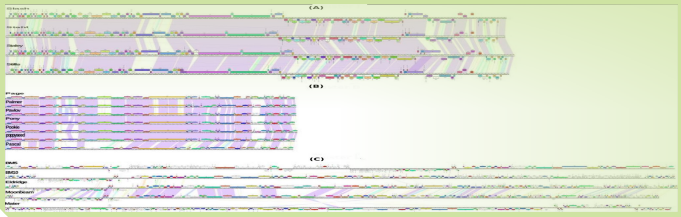

Fig.: Genomic organization of *Bacillus megaterium* phages. Phages mapped using Phamerator **(Richter, 2009)**; the purple lines between phages underline the regions with high similarity, while the ruler corresponds to genome base pairs. The predicted genes are presented as boxes either above or below the genome (line), depending on whether are rightwards or leftwards transcribed, respectively. Gene numbers are shown within each box; The phages maps showed by phams; genes colored according to their function categories "phams" While, the unfilled boxes refereed to the genes that show low similarity (orphams). *B. megaterium* phages: **A)** J cluster and **B)** D cluster show a high similarity between its members because all members belong to the same sub-clusters (D1 and J1), respectively. while **C)** L cluster phage members belong to different sub-clusters show a low similarity.

References

Funding Source: