FUNCTIONAL RELATEDNESS OF ENERGY PRODUCING RHODOPSEUDOMONAS PALUSTRIS GENOMES

Shaneka Simmons, Hari H.P. Cohly and Raphael D. Isokpehi

Center for Bioinformatics & Computational Biology, Department of Biology, Jackson State University, Jackson MS 39217, USA

Abstract: We have observed that 82% (538 of 624) completely sequenced prokaryotic genomes contain genes that encode proteins that have the universal stress protein domain (Pfam00582). Furthermore, the following bacterial species: Geobacter metallireducens, Shewanella oneidensis MR-1, Rhodopsuedomonas palustris, Desulfovibrio vulgaris, Deinococcus radiodurans R1, Nitrosomonas europaea, and Clostridium thermocellum, have been sequenced by the Department of Energy’s Joint Genome Institute and are relevant to research on sustainable bioenergy. Among these bacteria, only Rhodopsuedomonas had the highest number of sequenced genomes as well as highest number of genes encoding proteins with the USP domain. The universal stress proteins provide biological cells with the ability to respond to environmental stresses such as nutrient starvation, drought, high salinity, extreme temperatures and exposure to toxic chemicals. Rhodopseudomonas are rod-shaped, gram-negative, purple nonsulfur phototrophic bacteria belonging to the alpha subclass of the Proteobacteria, commonly found in various types of marine environments and soils. These organisms can grow in both anaerobic and aerobic conditions and have a genetic makeup that allows their genes to be easily removed from their system. Rhodopsuedomonas palustris are metabolically versatile species that can convert atmospheric carbon dioxide into biomass, recycle aromatic polymers of lignin, produce hydrogen, and fix nitrogen. Their ability to adapt and live under various environmental constraints as well as biodegrade pollution to be used as biofuel, make them a model system for research on renewable energy from biological sources. The sequence of a 6th Rhodopseudomonas palustris strain, DX-1, is in production. Strain DX-1 can produce high power densities that allow it to generate bioelectricity from the biodegradation of organic and inorganic waste in low-internal-resistance microbial fuel cells. There are five completed sequenced genomes of R. palustris: BisA53, BisB5, BisB18, CGA009, and HaA2. The gene count of these R. palustris strains ranged from 4492 to 5016. We hypothesize that genomes with similar functional profiles will cluster together. The objective of this study was to determine the functional relatedness of finished genomes of Rhodopseudomonas palustris based on Cluster of Orthologous protein Groups. Hierarchical Clustering and Correlation Matrix were generated using tools on the Integrated Microbial Genome Database. Two major clusters were obtained. Strains HaA2, CGA009 and BisH53 clustered together with HaA2 sharing the same node. Strains BisA53 and BisB18 shared the same node. The maximum and minimal correlation values between genomes were 0.84 and 0.95 respectively. We intend to compare the expression of genes encoding universal stress protein in various stress conditions.

Acknowledgements: Mississippi NSF-EPSCoR “Innovations through Computational Sciences” Award (EPS-0556308); Research Centers in Minority Institutions (RCMI) – Center for Environmental Health (NIH-NCRR G12RR13459-09).