PROTEIN LENGTH VARIATION IN UNIVERSAL STRESS PROTEINS FROM SALMONELLA GENOMES

Christina Bernard, Kafui Edusei, Tranquility Gordon, Stedman Ashley, Kenneth Campbell, Chidoze Adindu, Wellington K. Ayensu, Hari H.P. Cohly and Raphael D. Isokpehi

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

Abstract: The genus Salmonella consist of Gram-negative facultative rod-shaped bacteria that intestinal tract in warm-blooded animals. Salmonella causes typhoid fever, paratyphoid fever, and foodborne illness. In the United States, the repeated multistate outbreak of Salmonella infections associated include Tomato (2006), Pet Food (2006), Banquet Brand Pot Pies (2007), Veggie Booty (July 2007), Peanut Butter (2007), Malt-O-Meal Rice/Wheat Cereals (2008), Cantaloupe (2008), Peanut Butter (2009), Jalapeno Peppers (2009), Alfalfa Sprouts (2009), Raw Alfalfa Sprouts (2010), Marie Callender's Cheesy Chicken and Rice single-serve frozen entrée (June 2010) and Eggs (August 2010). There is a continuous need to understand the intrinsic features that enable bacteria pathogens to survive a variety of food processing conditions. Genes encoding the universal stress protein (USP) domain are known to provide a variety of pathogenic bacteria with ability to adapt to changing growth and host conditions. Differences in protein length can be biologically signification and are due to different evolutionary events and ecological conditions. The Integrated Microbial Genome Database contains 39 annotated Salmonella genomes encoding 236 genes encoding proteins with the universal stress protein domain. The objective of this study was to compare the protein length of proteins with the universal stress protein domain. A total of 36 unique amino acid lengths that ranged from 48aa to 905aa were observed. Twenty-eight of the unique lengths had only one example. The other eight clusters of proteins were based on the following lengths 76, 97, 142, 144, 131, 315, 894 and 904. Furthermore, 139 USPs had length of 131aa to 142aa. We have generated multiple sequence alignments of these clusters to propose new hypothesis sequence variation in the universal stress proteins of Salmonella.

Acknowledgements: Mississippi NSF-EPSCoR “Modeling and Simulation of Complex Systems” Award (EPS-0903787); Pittsburgh Supercomputing Center's National Resource for Biomedical Supercomputing (T36 GM008789); U.S. Department of Homeland Security Science & Technology Directorate (2007-ST-104-000007; 2009-ST-062-000014; 2009-ST-104-000021); Empowering Historically Underrepresented Undergraduate Students in Functional Genomics (NSF-DBI- 0958179) and Research Centers in Minority Institutions (RCMI) – Center for Environmental Health (NIH-NCRR G12RR13459). Disclaimer: “The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the funding agencies.