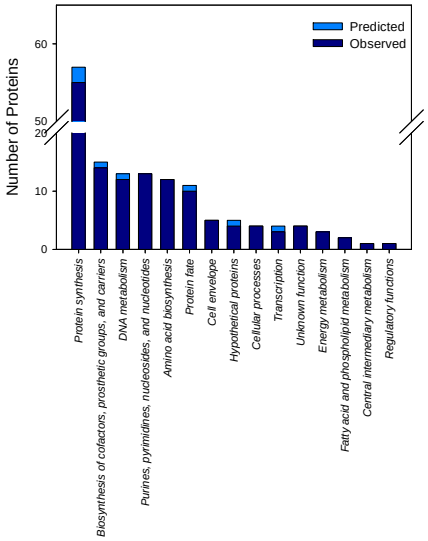
William R. Wiley
EMSL In Brief
Environmental Molecular Sciences Laboratory

Bacteria are Deeply Rooted Conservatives

Diverse bacteria have much in common, right down to the core. Using a peptide database collected at the Department of Energy's Environmental Molecular Sciences Laboratory—six years in the making, comprising approximately 967,000 peptides, and linked to protein and gene data—researchers from the Pacific Northwest National Laboratory discovered a group of proteins that are shared and commonly expressed by a wide variety of different bacteria. Called the "core proteome," the ubiquitous nature of this group of proteins suggests that they are fundamental to bacterial life. Knowing what proteins lie outside of the core proteome and thus they might be dispensable offers a new potential strategy for treating infections and creating designer bacteria for alternative energy production and bioremediation. The researchers analyzed 17 bacteria ranging from the ocean-dwelling *Pelagibacter ubique* to plague-causing *Yersinia pestis*. Their study revealed that the core proteome comprised 74% of the larger genome sequence-based core genome predictions. The core proteome appears to be responsible for the basal functions of bacteria; more than half is devoted to protein synthesis. Strikingly, about 7% of the core proteome is composed of proteins whose function is unknown. Because of the central nature of the core proteome, these proteins need further study. Proteins outside of the core proteome are responsible for the lifestyle-specific differences of individual bacterial species.



Functional categories (http://www.tigr.org/) assigned to the core genome and core proteome. The largest portion of the core proteome is involved in protein synthesis. However, several poorly characterized proteins were also observed as part of the core proteome, which highlights the need to better understand the function of these proteins.

Scientific Impact: Commonalities in the genome indicate shared ancestry, and proteomics verifies the expression of those genes. The team verified the expression of a core set of genes that defines bacterial life and designed an approach that can be used to compare the lifestyles of selected bacteria. This research builds on EMSL's work in predicting biological functions using high-throughput global-scale biomolecular studies.

Societal Impact: By identifying the core proteome, the identified non-core proteins dictate bacterial-specific differences by default. These specific proteins could potentially be used as environmental markers or therapeutic targets. In addition, understanding the core proteome will help researchers design synthetic organisms for environmental applications, such as alternative energy production and bioremediation.

Citation: Callister SJ, LA McCue, JE Turse, ME Monroe, KJ Auberry, RD Smith, JN Adkins, and MS Lipton. 2008. "Comparative Bacterial Proteomics: Analytics of the Core Genome Concept." *PLoS ONE* **3**(2):e1542.

Acknowledgements: Portions of this research were supported by the DOE Office of Biological and Environmental Research, and the National Institute of Health's National Institute of Allergy and Infectious Diseases and National Center for Research Resources.

P.O. Box 999 Richland, WA 99352 • http://www.emsl.pnl.gov • 509-376-2553

Pacific Northwest National Laboratory
Operated by Battelle for the U.S. Department of Energy
Office of Science
U.S. DEPARTMENT OF ENERGY