

ScholarWorks@UMass Amherst

Off-campus UMass Amherst users: To download dissertations, please use the following link to [log into our proxy server](#) with your UMass Amherst user name and password.

Non-UMass Amherst users, please click the view more button below to purchase a copy of this dissertation from Proquest.

(Some titles may also be available free of charge in our [Open Access Dissertation Collection](#) , so please check there first.)

Understanding genome structure, function, and evolution in the halophilic archaeon *Halobacterium* NRC-1

Sean P Kennedy, *University of Massachusetts Amherst*

Abstract

The genome of *Halobacterium* NRC-1 is 2,571,010 bp and encodes 2,630 proteins. The genome is arranged into 3 replicons, a large chromosome and two smaller mini-chromosomes, pNRC100 and pNRC200. The mini-chromosomes share approximately 145 Kb of sequence similarity and harbor the majority of the 91 IS elements. This large amount of repeated DNA required mapping pNRC200 for final assembly of the genome sequence. ^ The proteome of NRC-1 is extremely acidic, with an average pI of 4.9, and represents the major adaptation to the high internal salt concentration. Homology modeled structures showed the majority of additional charge is at the surface of the protein where it compensates for lower water activity and depressed dielectric constant to maintain solubility at saturating salt concentrations. GC and amino acid compositions reflect the extremely acid proteome. An overrepresentation of codons GAC and GAG, coding for aspartic acid and glutamic acid respectively, are found in the genome. ^ NRC-1 showed extensive evidence of lateral gene transfer (LGT). Although 16S rDNA analysis shows halophiles to be a monophyletic group branching from the methanogens in the Euryarchaeota, several features important to environmental adaptation appear to be acquired laterally from bacteria. The *nuo* and *men* genes corresponding to electron transport were identified as laterally transferred, giving evidence of the evolutionary past of this halophile. Further analysis of local databases and the Clusters of Orthologous Genes (COG) databases revealed, among the 2411 non-redundant protein coding genes from NRC-1, 783 proteins with bacterial character by BLAST score. Phylogenetic analysis showed several biosynthetic, transport, and energy systems, including histidine utilization, purine metabolism, and glycerol utilization, were likely acquired by LGT. However, no link between the 91 IS elements and the laterally transferred genes was found. ^ While the structure of the genome, with its three replicons, mini-chromosomes, IS elements, different GC-fractions, and observed plasticity, have been the driving force towards sequencing and subsequent analysis, the ultimate role of these various features remains mostly in the realm of the unknown. I discuss several scenarios to explain the presence of IS elements, duplication of genes, and overall genomic organization. ^

Subject Area

Genetics|Microbiology

Recommended Citation

Kennedy, Sean P, "Understanding genome structure, function, and evolution in the halophilic archaeon Halobacterium NRC-1" (2005). *Doctoral Dissertations Available from Proquest*. AAI3163680.
<https://scholarworks.umass.edu/dissertations/AAI3163680>

[View More](#)

DOWNLOADS

Since July 19, 2006

Share

COinS