An Analysis of Bipyrimidin Limitations as Photoprotective Genome Strategies in Halophilic Archaea

Daniel L. Jones and Bonnie K. Baxter, Ph.D.

Abstract

Halophiles experience high levels of ultraviolet (UV) light in their environments and demonstrate resistance to UV irradiation under laboratory conditions, yet the mechanisms underlying halophore photoprotection remain unclear. Herein, we consider genomic signatures as a potential photoprotective strategy. One of the predominant forms of UV-induced DNA damage is cyclobutane pyrimidine dimer (CPD) formation at pyrimidine dinucleotides, particularly at thymine-thymine (TT) and thymine-cytosine (TC) sites. A survey of all haloarchaea genomes available in the NCBI database demonstrates that on average, haloarchaea feature higher proportions of guanine (G) and C nucleotides within their genomes than non-halophilic microorganisms. Altogether, these notions have led us, among others, to consider whether the high G+C content seen in haloarchaea serves a photoprotective function through limiting T nucleotides. We reason that if halophiles have evolved a genomic strategy to attenuate the damaging effects of UV radiation, then the incidences of the most photoreactive dinucleotides should be limited in their genomes. Therefore, we designed a program to determine the frequencies of the four pyrimidine dinucleotides (TT, TC, CT, and CC) for samples of haloarchaea, archaea, and bacteria (n=29 each). The outputs were used to generate a single metric quantifying the “genomic photoreactivity” of each sample; we then employed statistical methods to compare results between the three sample groups. Our findings do not support the notion that the UV resistance seen in haloarchaea can be attributed to a genomic strategy.

Background

The harmful effects of UV radiation are largely attributed to the damage it causes to DNA. One of the primary forms of such damage is CPD formation between adjacent T and/or C nucleotides, which causes “kinks” in the affected DNA strand that may in turn give rise to mutations during replication. TT and TC are the most photoreactive sequences; therefore, a smaller incidence of TT and TC dinucleotides within a genome corresponds to a smaller propensity for UV-induced DNA damage.

Halophiles, the “salt-loving” microorganisms found ubiquitously in hypersaline environments such as Great Salt Lake, Utah, experience high levels of solar UV radiation and exhibit resistance to its effects under simulated conditions. Further, they typically possess a larger proportion of G and C nucleotides within their DNA than other microorganisms. Herein, we investigate whether this reduction in T nucleotides contributes to their observed UV resistance through evaluating the incidences of the photoreactive pyrimidine dinucleotides in their genomes and comparing them to that of non-halophilic microorganisms.

Methods

1. Comparing Mean G+C Contents

   1. To establish whether haloarchaea do possess elevated G+C content relative to other microorganisms, all available prokaryotic G+C contents in the NCBI database were sampled (33 haloarchaea and 5074 others).

   2. The mean G+C contents of haloarchaea vs. other prokaryotes were compared via an independent samples t-test.

2. Computing Pyrimidine Dinucleotide Incidences

   1. To determine the incidence of each of pyrimidine dinucleotide (TT, TC, CT, CC) appearing in a given DNA sequence, a dinucleotide-counting program was coded in R.

   2. Full genome sequences for all available haloarchaea and random samples of non-halophilic archaea and bacteria (n=29 each) were obtained from the NCBI database.

   3. The frequencies of each of the 4 pyrimidine dinucleotides and total number of nucleotides in the genomic DNA were counted for each sampled individual.

   4. The incidence of each dinucleotide was then established by dividing the number of dinucleotides counted by the number of nucleotides in the genome.

3. Assessing “Genomic Photoreactivity”

   1. To assess whether the genomic composition of haloarchaea is photoprotective, a metric quantifying “genomic photoreactivity” (Gp) was devised.

   2. For any given genome, Gp was determined by the incidences of the four pyrimidine dinucleotides, weighted by their intrinsic photoreactivities (Matallana-Surget et al., 2008):

\[
Gp = 1.73(\text{TCs}) + 1.19(\text{TTs}) + 0.61(\text{CTs}) + 0.39(\text{CCs})
\]

   3. Intergroup differences in Gp were compared between the three sample groups via a one-way ANOVA test and post-hoc Tukey contrasts.

Results & Conclusions

1. The results shown in Figure 2 demonstrate that on average, haloarchaea have higher G+C content than non-halophiles.

2. The results shown in Figure 3 indicate that haloarchaea have fewer incidences of TT & CT dinucleotides, yet higher incidences of TC & CC dinucleotides, than non-halophiles.

3. The results shown in Figure 4 give evidence that on average, the overall “genomic photoreactivity” (Gp) of haloarchaea genomes is smaller than that of archaea, but not bacteria.

References


Acknowledgments

This project was funded by a NASA Space Grant research infrastructure competitive award. We would also like to thank contributors to the NCBI genome database.

Figure 1. UV radiation damages DNA via inducing CPD formation between adjacent pyrimidine nucleotides, subsequently causing “kinks” in the DNA. (Image courtesy of: NASA/David Herrin)

Figure 2. G+C content (%) distributions for samples of haloarchaea (n=33) and other prokaryotes (n=5074). Means are indicated with ±1.96 standard errors. ***p < 1e-6.

Figure 3. Mean pyrimidine dinucleotide incidences for samples of haloarchaea and non-halophilic archaea & bacteria (n=29 each). Error bars represent ±1.96 standard errors.

Figure 4. Distributions for the computed “genomic photoreactivities” (Gp) of samples of haloarchaea & non-halophilic bacteria and archaea (n=29 each). Sample means are indicated with ±1.96 standard errors. ***p < 1e-6.