Bipyrimidine Signatures as a Photoprotective Genome Strategy in G+C-rich Halophilic Archaea

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

Background

Halophilic Archaea

- Experience high levels of ultraviolet (UV) radiation in their environments
- Demonstrate high resistance to UV
- Are protected by pigmentation and efficient DNA repair
- Have high genomic G+C content

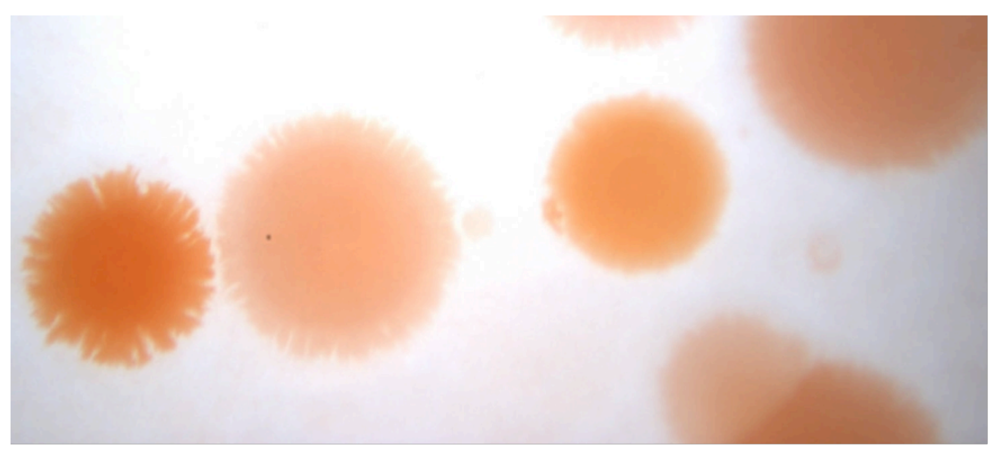

Figure 1. Halophilic archaea colonies from Great Salt Lake, Utah growing on salt agar

UV-induced DNA Damage

- The predominant forms of UV-induced DNA damage are cyclobutane pyrimidine dimers (CPDs)
- These form between adjacent pyrimidines
- Bipyrimidine photoreactivity is in the descending order of: TC > TT > CT > CC
- Limiting of the most photoreactive sequences should reduce overall genomic photoreactivity

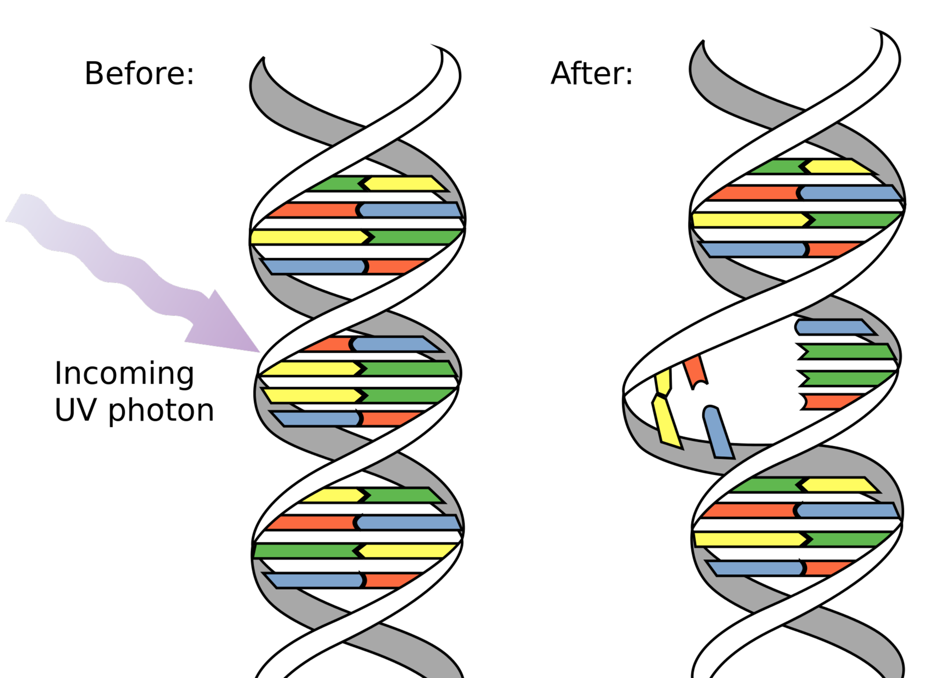

Figure 2. UV radiation damages DNA via inducing CPD formation between adjacent pyrimidine nucleotides, subsequently causing "kinks" in the DNA (Image courtesy of: NASA/David Herring)

Overarching Questions

- Do halophilic archaea have a net-photoprotective bipyrimidine signature?
- If so, how is it related to G+C content?
- Are photoprotective bipyrimidine signatures present among other taxa that live in high UV?

Results

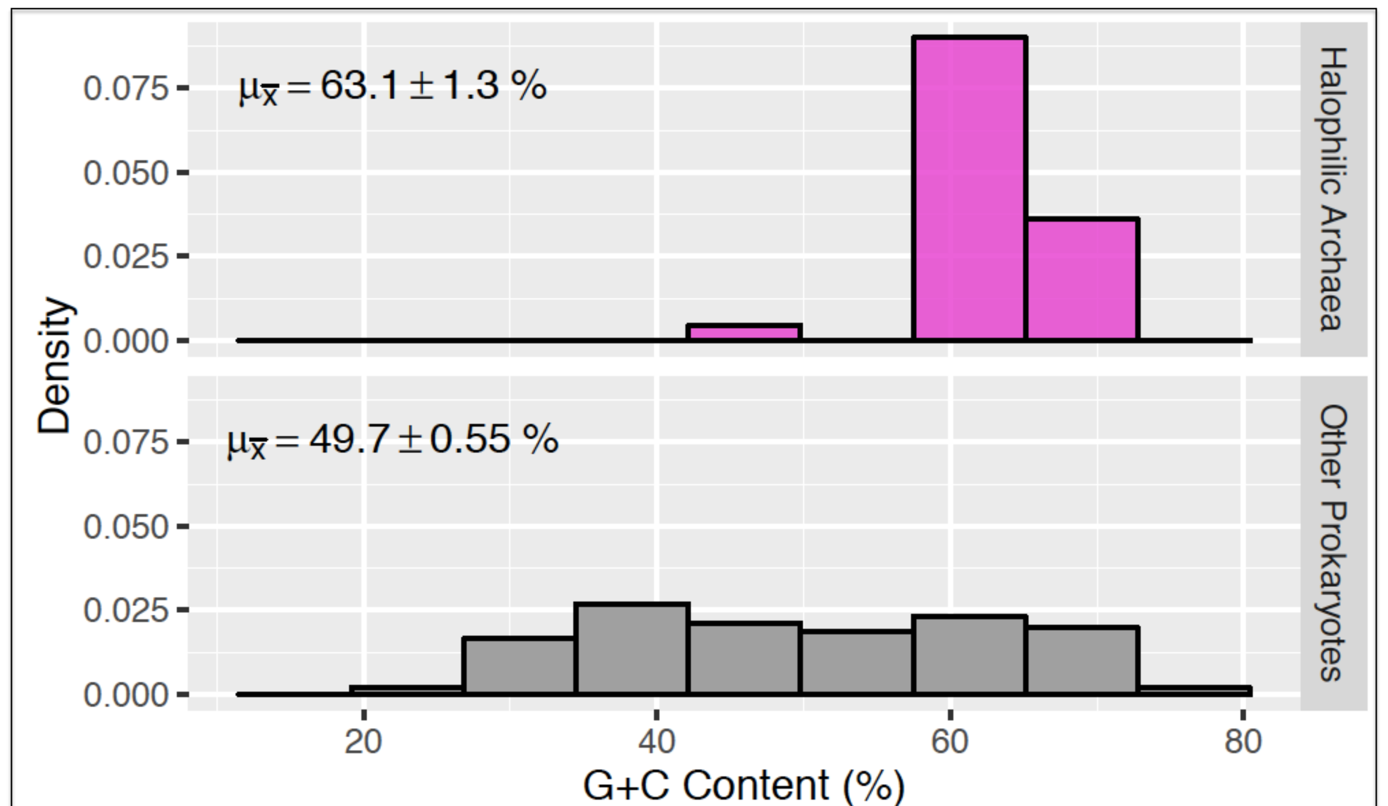

Figure 3. Genomic G+C content (%) distributions for samples of halophilic archaea (n = 29) and other prokaryotes (n = 2231). Sample means are denoted with +/− 1.96 standard errors. $p < 2.2 \times 10^{-16}$ (Welch Two Sample t-test)

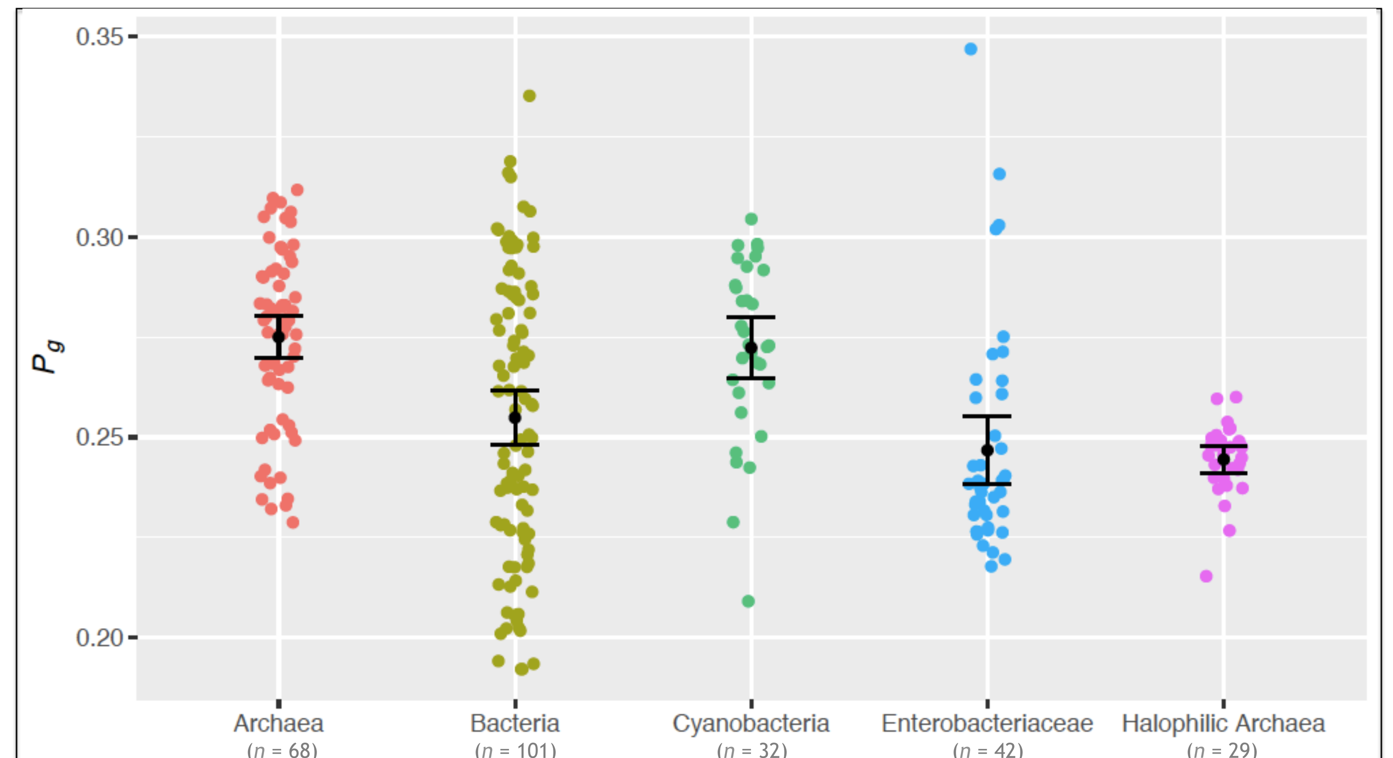

Figure 5. P_g distributions for each sample group. Error bars represent sample means +/− 1.96 standard errors. Intergroup differences were assessed via one-way ANOVA and post-hoc Tukey contrasts. Halophilic archaea and enterobacteriaceae have significantly smaller P_g than archaea and cyanobacteria ($p < 10^{-4}$).

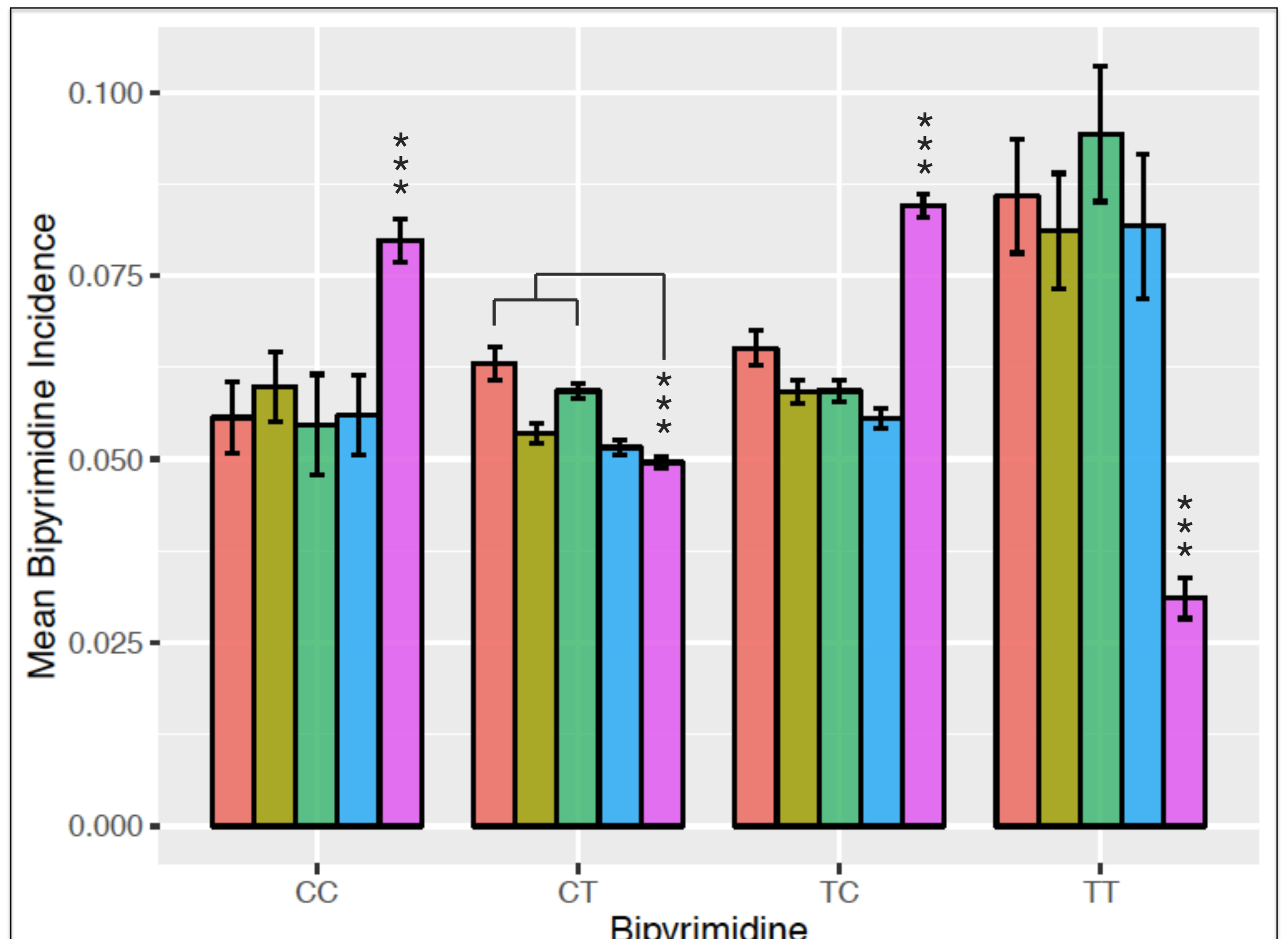

Figure 4. Mean bipyrimidine incidences for each sample group. Error bars represent +/− 1.96 standard errors. Intergroup differences were assessed via one-way ANOVA and post-hoc Tukey contrasts. Only differences pertaining to halophilic archaea are indicated. $***p < 10^{-4}$

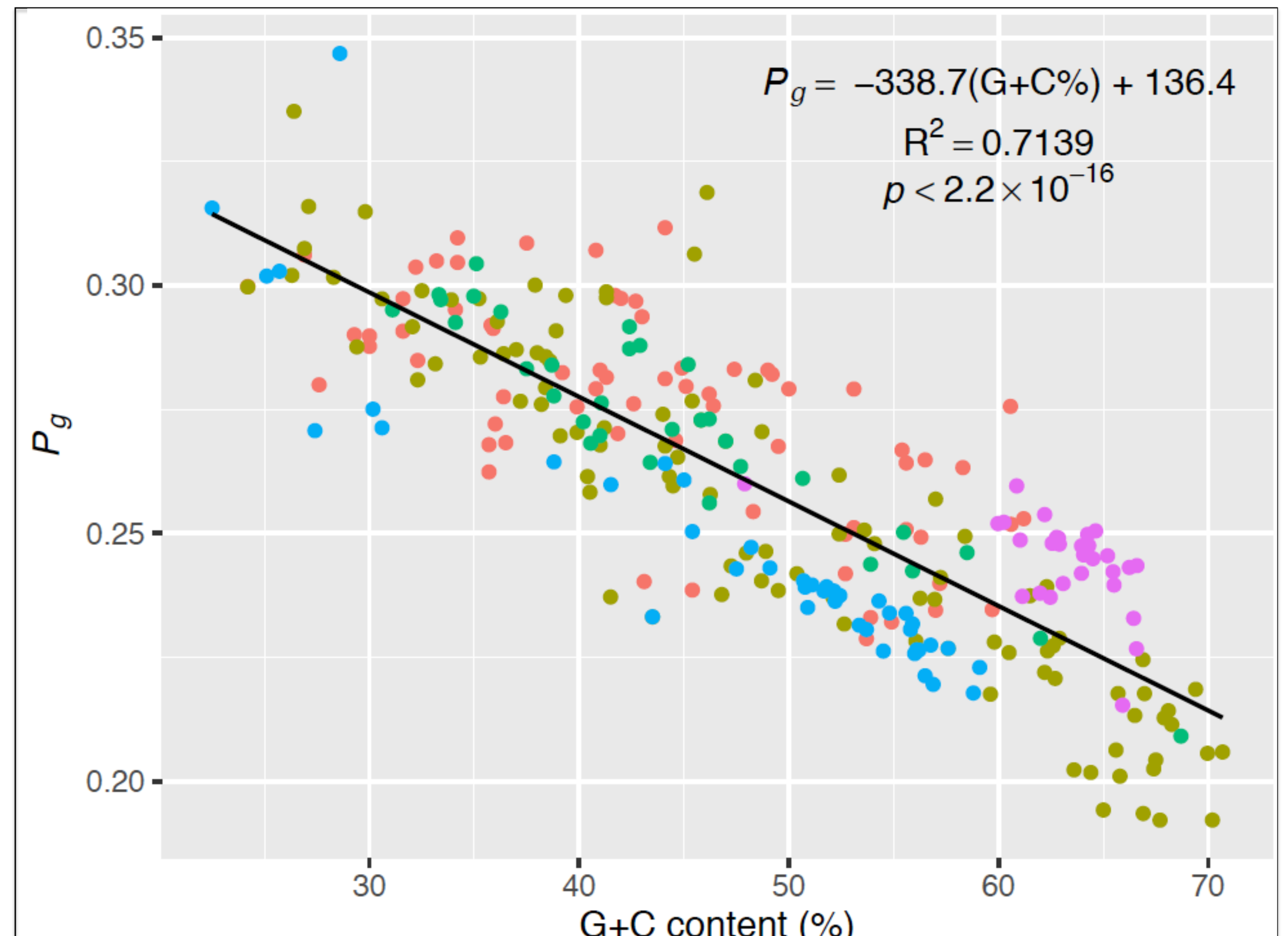

Figure 6. P_g versus G+C content (%) of each sampled genome (n_{total} = 272), with taxonomic group indicated by color. Regression analysis was conducted using a Pearson's product-moment correlation test.

Conclusions

1. There is a strong, negative correlation between P_g and G+C content (Figure 6)
   - This may be explained by the fact that the most photoreactive sequences are T-containing
2. We found no evidence that UV exposure is a selective pressure for low P_g
   - Enterobacteriaceae have similar P_g to halophilic archaea
   - Cyanobacteria have significantly higher P_g than both
3. The UV-resistance observed in halophilic archaea can be attributed in part to a genomic strategy

Methods

Genome Sampling

- Sequences were obtained from the NCBI database
- Four our G+C content analysis, one representative genome for each prokaryotic species presently available was sampled at random
- For all other analyses, we randomly sampled 1 halophilic archaea strain per species, 1 (non-halophilic) archaea, cyanobacteria, and enterobacteriaceae strain per genus, and 101 bacterial strains of unique genus

Determining Bipyrimidine Incidences

- We wrote a word-counting script in R to determine bipyrimidine frequencies within sampled genomes
- Bipyrimidine incidences (TC_i, TT_i, CT_i, CC_i) were computed by dividing frequency by genome size in bases

Determining Theoretical Genomic Photoreactivity (P_g)

- P_g corresponds to the weighted sum of a genome's bipyrimidine incidences:

$$P_g = 1.73(TC_i) + 1.19(TT_i) + 0.61(CT_i) + 0.39(CC_i)$$

- The weighting coefficients represent the intrinsic photoreactivity of each bipyrimidine, as determined by Matallana-Surget et al. (2008)

References

Baxter BK, *et al.* (2007) Great Salt Lake halophilic microorganisms as models for astrobiology: evidence for desiccation tolerance and ultraviolet radiation resistance. *Proc SPIE*. 6694.

Douki T, Cadet J. (2001) Individual determination of the yield of the main UV-induced dimeric pyrimidine photoproducts in DNA suggests a high mutagenicity of CC photolesions. *Biochemistry*. 30: 2495-2501.

Jones DL, Baxter BK. (2016) Bipyrimidine signatures as a photoprotective genome strategy in G + C-rich halophilic archaea. *Life*. 6: 37.

Matallana-Surget S, *et al.* (2008) Effect of the GC content of DNA on the distribution of UVB-induced bipyrimidine photoproducts. *Photochem & Photobiol Sci*. 7: 794-801.

Acknowledgments

We would like to thank the Utah NASA Space Grant Consortium and the Lawrence T. Dee-Janet T. Dee Foundation for funding. In addition, we express much gratitude to Jaimi Butler, Kendall Tate, and Chrono Nu for their assistance in the development of this project.