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

Daniel L. Jones and Bonnie K. Baxter *

Great Salt Lake Institute and Biology Program, Westminster College, Salt Lake City, UT 84105, USA; dlj0317@westminstercollege.edu
* Correspondence: bbaxter@westminstercollege.edu; Tel.: +1-801-832-2345

Abstract: Halophilic archaea experience high levels of ultraviolet (UV) light in their environments and demonstrate resistance to UV irradiation. DNA repair systems and carotenoids provide UV protection but do not account for the high resistance observed. Herein, we consider genomic signatures as an additional photoprotective strategy. The predominant forms of UV-induced DNA damage are cyclobutane pyrimidine dimers, most notoriously thymine dimers (TˆTs), which form at adjacent Ts. We tested whether the high G + C content seen in halophilic archaea serves a photoprotective function through limiting T nucleotides, and thus TˆT lesions. However, this speculation overlooks the other bipyrimidine sequences, all of which capable of forming photolesions to varying degrees. Therefore, we designed a program to determine the frequencies of the four bipyrimidine pairs (5′ to 3′: TT, TC, CT, and CC) within genomes of halophilic archaea and four other randomized sample groups for comparison. The outputs for each sampled genome were weighted by the intrinsic photoreactivities of each dinucleotide pair. Statistical methods were employed to investigate intergroup differences. Our findings indicate that the UV-resistance seen in halophilic archaea can be attributed in part to a genomic strategy: high G + C content and the resulting bipyrimidine signature reduces the genomic photoreactivity.

Keywords: halophilic archaea; DNA damage; DNA repair; photoprotection; cyclobutane pyrimidine dimer; bipyrimidine; genome evolution

1. Introduction

Halophilic microorganisms thrive in the briny waters of salt lakes and evaporative solar salterns [1]. Saturated salinity, from 25%–35%, presents osmotic obstacles for life, but we see representatives from all three domains thriving in these conditions [2]. Hypersaline habitats select for archaea at a higher abundance than bacteria and eukarya in the community [3]. The present study focuses on this particular group, which likely share genetic strategies for life at high salinity [2]. Microbial diversity studies show a wide array of halophilic archaea present in hypersaline ecosystems [4] that are adapted to thrive in the saltiest places on Earth.

Beyond salt, high solar radiation is a feature of these environments; thus, their microbial inhabitants must have evolved to overcome the challenge of ultraviolet (UV) exposure. Halophilic archaea are highly resistant to UV light, first noted by Dundas and Larsen [5]. For example, one *Halorubrum* species was previously found to be nearly ten-fold more UV-resistant than *Escherichia coli* [6]. This is due in part to their robust UV DNA repair systems, including photoreactivation and nucleotide excision repair [7–11], but there are clearly additional processes that afford the high level of resistance to solar radiation exhibited by halophilic archaea.

Several studies have proposed other potential photoprotective strategies; for example, carotenoids in the cell membranes of halophilic archaea may provide UV resistance [6,12,13], although the

mechanism of which remains unclear. These organisms have a unique composition of carotenoids that set halophilic archaea apart [14–17]. In fact, some halophiles may have an array of genes that allow their complex interplay with light [18], including carotenoid biosynthetic pathways, and also genes encoding retinal-containing proteins like bacteriorhodopsin, or "purple membrane" protein, which pumps protons out of the cell upon exposure to light [19] and may aid in the generation of ATP. An environmental genomics approach to saltern communities found the presence of a large variation of rhodopsin-like genes [20], indicating the significance of photobiology for halophiles.

Table 1. Halophilic archaea sampled for the present study, representative of all species with full genome sequences presently available (taxid: 183963). Data were obtained from the National Center for Biotechnology Information (NCBI) database [21].

Species	Strain	G + C (%)	Size (Mb)	Genes
Halalkalicoccus jeotgali	B3	62.56	3.699	3717
Halanaeroarchaeum sulfurireducens	HSR2	62.86	2.210	2213
Haloarcula hispanica	N601	62.47	3.902	3825
Haloarcula marismortui	ATCC 43049	61.14	4.275	4226
Haloarcula sp.	CBA1115	61.98	4.225	4108
Halobacterium hubeiense	JI20-1	66.58	3.130	3189
Halobacterium salinarum	NRC-1	65.92	2.571	2629
Halobacterium sp.	DL1	66.44	3.163	3237
Haloferax gibbonsii	ARA6	66.24	3.918	3783
Haloferax mediterranei	ATCC 33500	60.26	3.905	3800
Haloferax volcanii	DS2	65.46	4.013	3925
Halogeometricum borinquense	DSM 11551	59.97	3.944	3838
Halomicrobium mukohataei	DSM 12286	65.51	3.332	3293
Halopiger xanaduensis	SH-6	65.20	4.355	4174
Haloquadratum walsbyi	DSM 16790	47.90	3.179	2827
Halorhabdus tiamatea	SARL4B	62.77	3.146	3069
Halorhabdus utahensis	DSM 12940	62.90	3.117	2969
Halorubrum lacusprofundi	ATCC 49239	63.95	3.693	3523
Halostagnicola larsenii	XH-48	60.87	4.131	3954
Haloterrigena turkmenica	DSM 5511	64.25	5.441	5074
Halovivax ruber	XH-70	64.30	3.224	3187
Natrialba magadii	ATCC 43099	61.03	4.444	4128
Natrinema pellirubrum	DSM 15624	63.96	4.354	4178
Natrinema sp.	J7-2	64.06	3.794	3681
Natronobacterium gregoryi	SP2	62.20	3.788	3720
Natronococcus occultus	SP4	64.63	4.314	4162
Natronomonas moolapensis	8.8.11	64.50	2.913	2793
Natronomonas pharaonis	DSM 2160	63.08	2.750	2799
Salinarchaeum sp.	Harcht-Bsk1	66.60	3.255	3036

One photoprotective mechanism suggested in the literature for decades [22–26] relates to the genomes of organisms with high Guanine+Cytosine (G + C) content, such as halophilic archaea. Their genomes typically exceed 60% G + C (Table 1), which effectively reduces the number of thymines (Ts) present. Thymine dimers (TˆT), which form from the UV-induced cyclization of two adjacent thymines on the same DNA strand (Figure 1), were long thought to be the primary DNA damage arising from sunlight [23]. Logically, the limitation of T residues is expected to reduce the incidence of adjacent Ts on the same DNA strand and thus, the possibility of TˆTs. Reduction of DNA damage would reduce mutagenesis during replication of unrepaired lesions. Kennedy et al. [26] suggested that this strategy of thymine limitation could explain the UV resistance of halophilic archaea, but the idea remains untested.
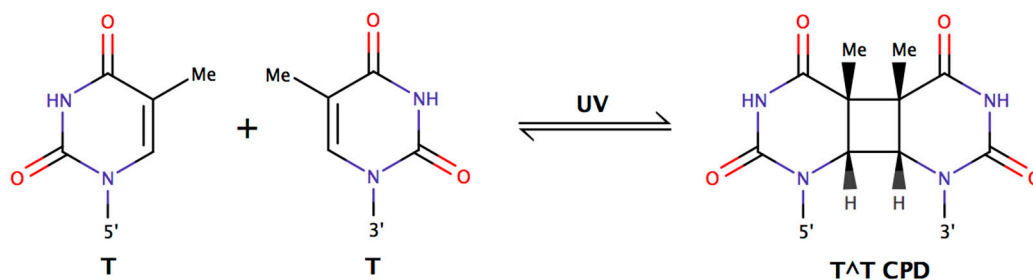
Figure 1. Thymine-thymine cyclobutane pyrimidine dimer (TˆT CPD) forms via butane ring cyclization between adjacent thymines on the same strand of DNA. Similar chemistry occurs at the other bipyrimidine nucleotides TC, CT, and CC [27].

The most important oversight in this premise is that TˆT lesions are but a subclass of cyclobutane pyrimidine dimers (CPDs), which include each possible adjacent bipyrimidine pair: (5′ to 3′) TˆT, CˆC, CˆT, and TˆC. Haynes [21] noted that the sensitivity of microorganisms to UV light was not mathematically proportional to the thymine frequency in the genome, which pointed to other lethal lesions. Any set of two adjacent pyrimidines may cyclize upon exposure to a photon of UV light. Though CPDs represent the majority of solar-induced DNA damage events [28], other potential lesions resulting from the UV-irradiation of bipyrimidine sequences are pyrimidine (6-4) pyrimidone photoproducts (6-4PPs). The proportion of CPDs to 6-4PPs is dependent on wavelength of light [29] and on flanking sequences [30,31].

Though it is logical to assume TˆT lesions may be limited in halophilic archaea, other photolesions may form; therefore, the photoreactivities of each bipyrimidine pair and the relationship between G + C content and their incidences must be considered in order to assess any photoprotective benefit [32,33]. Therefore, we sought to address the hypothesis that thymine limitation is photoprotective using a genome photoreactivity score that incorporates all bipyrimidine pairs and their respective photoreactivities, while accounting for all potential types of UV damage (e.g. CPDs and 6-4PPs). We developed an equation to quantify the theoretical photoreactivity of a genome based on the frequencies of each bipyrimidine doublet within it, weighted by their intrinsic photoreactivities (as determined by [33]) and genome size. Comparison with other groups and a robust statistical analysis sheds light on an interesting story in genome evolution.

2. Materials and Methods

2.1. Comparing G + C Content of Halophilic Archaea Versus Other Prokaryotes

A list of all prokaryotes with full-genome sequences available (*n* = 5074) and their corresponding G + C contents was downloaded from the NCBI database [21]. One representative genome for each species was selected at random, yielding sample groups of *n* = 29 halophilic archaea and *n* = 2231 other prokaryotes.

2.2. Genome Sampling

Tabulated lists of all species with full-genome sequences available were downloaded from the NCBI database [21] for the sample groups bacteria (taxid: 2, *n* = 4829), halophilic archaea (taxid: 183963, *n* = 33), archaea excluding halophilic archaea (taxid: 2157, *n* = 209), cyanobacteria (taxid: 1117, *n* = 90) and enterobacteriaceae (taxid: 91347, *n* = 736).

Steps were taken to minimize sample bias: for the halophilic archaea group, one representative strain from each species was selected at random, yielding a sample group of *n* = 29 halophilic archaea. For the archaea, cyanobacteria, and enterobacteriaceae groups, one representative from each genus was selected at random, yielding sample groups of *n* = 68 (non-halophilic) archaea, *n* = 32 cyanobacteria,

and *n* = 42 enterobacteriaceae. For the bacteria, the first 101 strains of a unique genus to be randomly selected constituted the final sample group of *n* = 101 bacteria.

The full-genome sequences corresponding to each sampled strain were downloaded as .fasta files from the NCBI database [21].

2.3. Determining Bipyrimidine Incidences

To determine the incidences (i.e., relative frequencies) of each bipyrimidine in the sampled genomes, a novel word-counting program "DinucleotideCounts" was written in the scripting language R (script available at: [34]). This program determines the frequency of each dinucleotide, the frequency of each nucleotide, and the size of any .fasta-formatted DNA sequence.

The bipyrimidine frequencies within sampled genomes were determined via the DinicleotideCounts program. Bipyrimidine incidences (TC_i, TT_i, CT_i, CC_i) were then computed by dividing each bipyrimidine's frequency by the size of the corresponding genome in bases.

*2.4. Determining Theoretical Genomic Photoreactivity (P_g)*

We devised the metric P_g to quantify the theoretical photoreactivity of a genome based on its bipyrimidine signature: P_g corresponds to the weighted sum of a genome's bipyrimidine incidences (TC_i, TT_i, CT_i, CC_i):

$$P_g = 1.73(TC_i) + 1.19(TT_i) + 0.61(CT_i) + 0.39(CC_i) \qquad (1)$$

The coefficients represent the intrinsic photoreactivity of each bipyrimidine, as experimentally determined by [33] via establishing the ratio between the frequency of the photoproducts (6-4PPs and CPDs) and the bipyrimidine incidences in DNA with varying G + C content.

2.5. Statistical Methods

G + C content averages were compared (Figure 2) via a Welch Two Sample t-test using the "stats" package in R [35]. Intergroup differences in bipyrimidine incidences (Figure 3) and P_g (Figure 4) were assessed via one-way analysis of variance (ANOVA) and post-hoc Tukey contrasts using the "multcomp" package [36] in R [35]. The regression analysis of G + C content vs. P_g (Figure 5) and corresponding Pearson's product-moment correlation test were carried out using the "stats" package in R. All randomization was facilitated by the "RAND" function in Microsoft Excel.
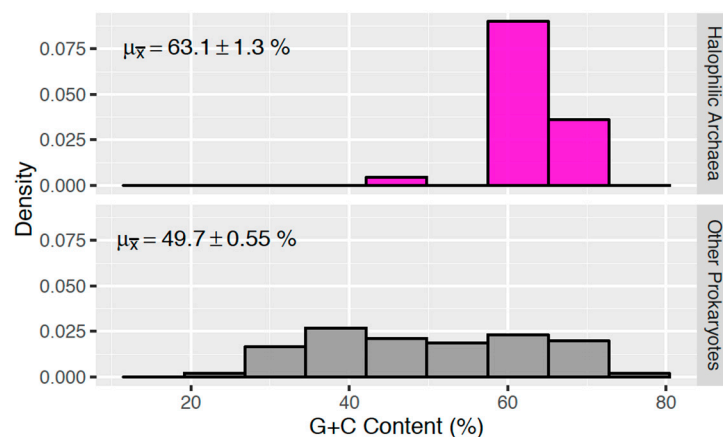

Figure 2. Genomic G + C content (%) distributions for samples of halophilic archaea (*n* = 29) and other prokaryotes (*n* = 2231). Sample means are denoted with +/− 1.96 standard errors. $p < 2.2 \times 10^{-16}$.
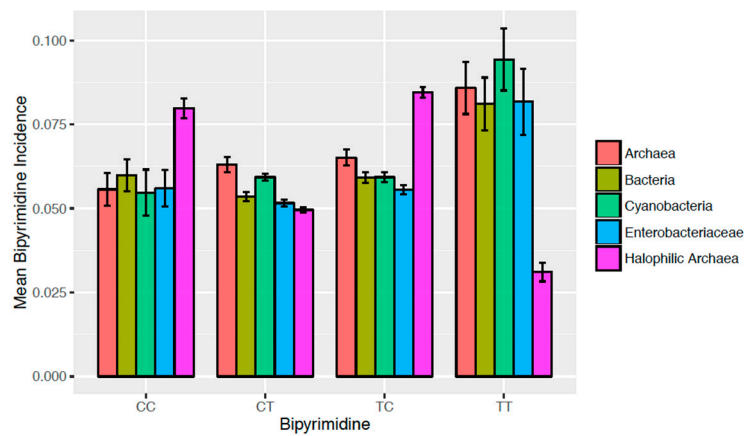
Figure 3. Mean bipyrimidine incidences for samples of (non-halophilic) archaea ($n = 68$), bacteria ($n = 101$), cyanobacteria ($n = 32$), enterobacteriaceae ($n = 42$), and halophilic archaea ($n = 29$). Error bars represent $+/-$ 1.96 standard errors.
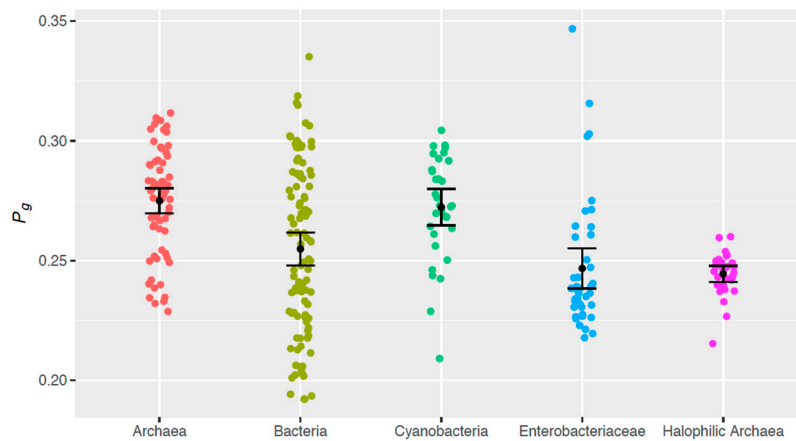
Figure 4. Distributions of the theoretical genomic photoreactivity P_g computed for samples of (non-halophilic) archaea ($n = 68$), bacteria ($n = 101$), cyanobacteria ($n = 32$), enterobacteriaceae ($n = 42$), and halophilic archaea ($n = 29$). Sample means are marked with error bars representing $+/-$ 1.96 standard errors.
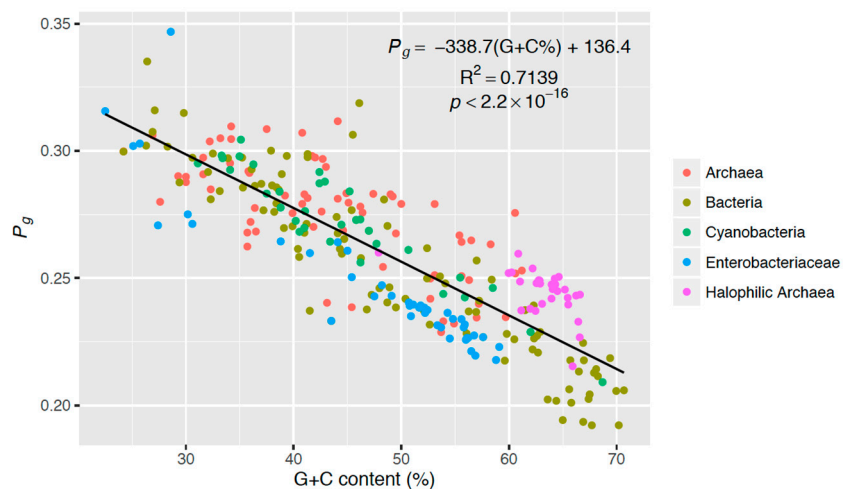
Figure 5. Theoretical genomic photoreactivity P_g versus G + C content (%) of each sampled genome ($n = 272$). The genomes of the five taxonomic groups are indicated by color.

3. Results

3.1. G + C Content of Halophilic Archaea

Halophilic archaea are known to have high G + C content with one notable exception, *Haloquadratum walsbyi* [37–39]. To understand the unique genomic features of these halophiles, we compare them to other prokaryotes. The NCBI Genbank database contains the full genome sequences for 29 species of halophilic archaea (Table 1), which were utilized for this study. 2231 other prokaryotic species (bacteria and non-halophilic archaea) were selected as described in Materials and Methods. This analysis (Figure 2) showed that halophilic archaea (excluding *H. walsbyi*) have a clustered distribution with a remarkably high G + C content (63.1% ± 1.3%) relative to other microbial life (49.7% ± 0.55%), thereby demonstrating the uniqueness of this group and pointing to the relatedness of halophilic archaea.

3.2. Bipyrimidine Signature of Halophilic Archaea

Due to their high G + C content (Figure 2), we predicted halophilic archaea would have low and high incidences of TT and CC dinucleotides, respectively. However, considering the incidences of all bipyrimidine pairs (CC_i, CT_i, TC_i, and TT_i) is necessary for understanding the overall photoreactivity of their genomes. Hence, bipyrimidine incidences were determined for and subsequently compared between halophilic archaea and the other taxonomic groups bacteria, non-halophilic archaea, cyanobacteria, and enterobacteriaceae (*as described in Materials and Methods*) for insight into unique genomic signatures. If an evolutionary genome strategy were employed to protect organisms from sunlight, one would expect that a photosynthetic group like cyanobacteria might share such a strategy, and that enterobacteriaceae bacteria, which dwell inside higher eukaryotes, would not. Random samplings of species in these groups are thus included as controls, in addition to the bacteria and non-halophilic archaea (Figure 3).

As predicted, halophilic archaea are distinctive among these control groups in their low TT_i and high CC_i (Figure 3). Further, it was found that halophilic archaeal genomes are also characterized by high TC_i relative to the comparative groups. Multiple comparisons of means with respect to each bipyrimidine pair (*as described in Materials and Methods*) point to some interesting observations, indicating that on average:

- Halophilic archaea have larger CCi than any other group ($p < 10^{-4}$ each). No other significant intergroup differences in CCi were detected.
- Halophilic archaea have smaller CTi than archaea and cyanobacteria ($p < 10^{-4}$ each). Other significant differences in CTi were found between enterobacteriaceae vs. archaea ($p < 10^{-4}$), cyanobacteria vs. bacteria ($p = 2.35 \times 10^{-4}$), and enterobacteriaceae vs. cyanobacteria ($p < 10^{-4}$).
- Halophilic archaea have larger TCi than any other group ($p < 10^{-4}$ each). Archaea have the next highest level of TCi, being larger than bacteria ($p < 10^{-4}$), cyanobacteria ($p = 3.27 \times 10^{-3}$), and enterobacteriaceae ($p < 10^{-4}$).
- Halophilic archaea have smaller TTi than any other group ($p < 10^{-4}$ each). No other significant intergroup differences in TTi were detected.

3.3. Intergroup Differences in Theoretical Genomic Photoreactivity (P_g)

Until recently, TTs were thought to be the most photoreactive sequences [24]. This idea has been challenged by current data on the photoreactivity of each bipyrimidine pair in both naked and intracellular DNA with a variety of genome G + C contents [31,32]. Matallana-Surget et al. [33] experimentally found the relative photoreactivities of the bipyrimidine sequences to be in the decreasing order of TC > TT > CT > CC. Furthermore, these authors quantified the intrinsic photoreactivity of each bipyrimidine via determining ratio between frequency of photoproducts (CPDs and 6-4PPs) and bipyrimidine incidences in DNA with varying G + C content. These ratios were

employed as coefficients in the equation for our theoretical quantification of genomic photoreactivity in terms of bipyrimidine signature, P_g, (Equation (1)). P_g was calculated for each sampled genome and then compared across all tested taxonomic groups (Figure 4).

Intergroup differences in P_g were examined via a multiple comparison of means (*as described in Materials and Methods*), which indicated that on average:

- Halophilic archaeal genomes have a smaller P_g than do genomes from (non-halophilic) archaea ($p < 10^{-3}$) and cyanobacteria ($p = 1.01 \times 10^{-3}$).
- Similarly, enterobacteriaceae genomes are less photoreactive than (non-halophilic) archaea ($p < 10^{-3}$) and cyanobacteria ($p = 1.02 \times 10^{-3}$).
- (Non-halophilic) archaeal genomes have a larger P_g than do genomes from the other sample groups, with the exception of cyanobacteria ($p < 10^{-3}$).

3.4. Genomic Strategy of Photoprotection

In analyzing photoreactivity in the studied genomes, we plotted each computed P_g value against its corresponding genome's G + C content (Figure 5). These data clearly show a strong, negative correlation between P_g and G + C content ($R^2 = 0.7139$, $p < 2.2 \times 10^{-16}$). Figure 5 further demonstrates that this relationship holds true not only for the halophilic archaea tested, but also for all other genomes with this bias toward G + C, altogether giving evidence that an organism's genomic G + C content contributes to their relative sensitivity to UV radiation.

4. Discussion

Solar exposure for all prokaryotic life on Earth results in UV-induced DNA damage, the majority of which is in the form of CPDs, a cycloaddition of two adjacent pyrimidine bases (Figure 1) [28,31]. CPDs are mitigated by two universal DNA repair systems: photolyase [40,41] and nucleotide excision repair [42]. Halophilic archaea are no exception as they have a similar ratio of CPD to 6, 4 photoproducts [8] and efficient repair [7–11].

The UV spectrum is broken into three subdivisions: UV-A (315–400 nm), UV-B (280–315 nm), and UV-C (<280 nm), with UV-C being the most damaging for DNA [43] due to the shorter wavelengths. The solar spectrum (UV A/B) peaks at 300 nm and results in CPDs as the principal type of UV-lesions, placing less emphasis on 6-4PPs [27]. UV-C light is sometimes used in laboratory experiments to amplify DNA damage detection [6,24], but it does not have real world consequence for biological systems since it is absorbed by the oxygen and ozone in the Earth's atmosphere [44].

The photoreactivity of a specific bipyrimidine pair may vary depending on the wavelength of UV light utilized, the flanking sequence conditions, and the cellular environment. For example, in one study the ratio of TˆT to other CPDs was greater when UV-C light was used in experiments than when UV-B was utilized [30]. This may account for the significance placed on the TˆT lesions over other photoproducts in decades of laboratory experimentation. The bipyrimidine photoreactivity coefficients utilized in our P_g calculations were adapted from a study that measured UV-B induced CPDs and 6-4PPs [33].

Our data clearly show that T limitation reduces TˆT formation (Figure 3). However, this is an incomplete picture. We developed the P_g formula (Equation 1) to further assess other bipyrimidine impacts on photoreactivity. For example, the corresponding enrichment in C nucleotides and the higher photoreactivity of the TC bipyrimidine [33] also impact the P_g score for a halophilic archaeal genome. Nevertheless, the strong, negative correlation observed between P_g and G + C content (Figure 5) gives evidence that T limitation facilitates a net increase in photoprotection. Note that the three most photoreactive sequences, TC, TT, and TC [32,33], are T-containing: this could explain the observed relationship between genomic photoreactivity and G + C content. It should also be noted that CˆC photolesions are the most mutagenic [32], which suggests that there is more work to be done beyond photoreactivity to explore mutagenesis from solar irradiation. Finally, determination of the

role that flanking sequences [30] play in the formation of photolesions could add another dimension to this work.

From a genome evolutionary perspective, if exposure to UV light was driving an organism's genomic bipyrimidine signature, we would expect to see this adaptation in cyanobacteria, a photosynthetic group of microorganisms, and we do not (Figure 4). Conversely, we would expect members of the enterobacteriaceae group, which dwell inside higher eukaryotes protected from the sun, to be void of UV adaptations, but in fact, their photoreactivity scores are more similar to halophilic archaea than any other sample group (Figure 4). The analysis of our controls gives evidence that there is not a clear/predictive relationship between P_g and UV-resistance alone. The similarity between the enterobacteriaceae and halophilic archaea stems not from lifestyle, but instead from their relativity high G + C contents (Table 2), which have been shown to result in low P_g values (Figure 5). Hence, we find no evidence in our analysis of genomic photoreactivity that UV exposure is a selective pressure for a photoprotective bipyrimidine signature.

Table 2. G + C contents (%) of all sample groups utilized in the present study (archaea taxid: 2157, *n* = 68; bacteria taxid: 2, *n* = 101, cyanobacteria taxid: 1117, *n* = 32; enterobacteriaceae taxid: 91347, *n* = 42; halophilic archaea taxid: 183963, *n* = 29). Sample means are denoted with +/− 1.96 standard errors. Data was obtained from the NCBI database [21].

Sample Group	Mean G + C Content (%)	Median (%)
(Non-Halophilic) Archaea	43.87 +/− 2.30	43.05
Bacteria	49.10 +/− 2.55	47.25
Cyanobacteria	44.21 +/− 3.08	42.65
Enterobacteriaceae	47.83 +/− 3.17	51.73
Halophilic Archaea	63.09 +/− 1.28	63.95

One important outlier is *Haloquadratum walsbyi*, a dominant halophilic archaeon in salt lakes and salterns. This microorganism is an exception to the G + C-rich genomes of the other members of this group, having only 48% G + C content (Table 1) [37–39]. This square-shaped organism otherwise shares the same ecosystem niche as the other extreme halophiles, thriving in high UV exposure; however, it lacks the photoprotective genome signature seen in other halophilic archaea, having a P_g value of 0.260, which is over 8 standard errors smaller than the sample group's mean of 0.245 (Figure 4). Bolhuis et al. [37] noted that *H. walsbyi* has a relatively higher number (four) photolyase genes, and this could impact its ability to counteract solar DNA damage.

If the overwhelming majority of halophilic archaea have genomes with lower photoreactivity because of their high G + C contents (Table 2), then we must ask how the G + C richness evolved? Litchfield [45] pointed to the stability of genomes rich in base pairs that have more hydrogen bonds; GC pairs have three and AT pairs have two. Halophilic archaea living in high salinity (and having high intracellular cation concentrations) would benefit from DNA helices that are more tightly paired, adding stability to their molecular structure in a destabilizing environment. High C + C content has also been discussed as a hypersaline adaptation that impacts the proteome composition of these microorganisms [46]. A + T limitations result in preferences for particular (G + C-rich) amino acids over others. For example, acidic residues such as aspartic acid would be over-represented and cysteine would be under-represented. Paul et al. demonstrate that this signature would reduce the likelihood of helices forming in protein tertiary structure but would positively impact coil structures. Both stronger DNA hydrogen bonding and preferential codon usage (and resulting protein structure) could impact the ability of halophilic archaea to manage life in hypersaline waters.

Genomic signatures, such as bipyrimidine limitations, should be examined when painting a picture of genome evolution for any microbial community. Each related group of species in their environment over time results in a unique signature of nucleic acid and protein composition [47–49], and G + C content is a marker for this signature. The UV-resistance observed in halophilic archaea can

certainly be attributed to a genomic strategy. However, the driver for this is not simply UV light as other evolutionary pressures, including hypersalinity, are also significant. Indeed, high G + C reduces the photoreactivity due to the specific bipyrimidine composition, but it also impacts a combination of factors including DNA, RNA and protein-level features. Life for halophilic archaea must be well-suited for the dual extremes of both UV exposure and hypersalinity.

References

1. Baxter, B.K.; Litchfield, C.D.; Sowers, K.; Griffith, J.D.; DasSarma, P.A.; DasSarma, S. Great salt lake microbial diversity. In *Adaptation to Life in High Salt Concentrations in Archaea Bacteria, and Eukarya*; Gunde-Cimerron, N., Oren, A., Plemenita, A., Eds.; Springer: Dordrecht, The Netherlands, 2005; pp. 9–21.

2. Oren, A. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Syst.* **2008**, *4*, 2. [CrossRef] [PubMed]

3. Grant, W.D.; Gemmell, R.T.; McGenity, T.J. Halophiles. In *Extremophiles: Microbial Life in Extreme Environments*; Horikoshi, K., Grant, W.D., Eds.; Wiley-Liss: New York, NY, USA, 1998; pp. 93–132.

4. Almeida-Dalmet, S.; Sikaroodi, M.; Gillevet, P.M.; Litchfield, C.D.; Baxter, B.K. Temporal study of the microbial diversity of the north arm of Great Salt Lake, Utah. *Microorganisms* **2015**, *3*, 310–326. [CrossRef]

5. Dundas, I.D.; Larsen, H. A study on the killing by light of photosensitized cells of *Halobacterium Salinarium*. *Arch. Microbiol.* **1963**, *46*, 19–28. [CrossRef]

6. Baxter, B.K.; Eddington, B.; Riddle, M.R.; Webster, T.N.; Avery, B.J. Great Salt Lake halophilic microorganisms as models for astrobiology: Evidence for desiccation tolerance and ultraviolet radiation resistance. In Proceeding of the SPIE 6694, Instruments, Methods, and Missions for Astrobiology X, San Diego, CA, USA, 26 August 2007; p. 669415. [CrossRef]

7. Hescox, M.A.; Carlberg, D.M. Photoreactivation in *Halobacterium cutirubrum*. *Can. J. Microbiol.* **1972**, *18*, 981–985. [CrossRef] [PubMed]

8. McCready, S. The repair of ultraviolet light-induced DNA damage in the halophilic archaebacteria, *Halobacterium cutirubrum, Halobacterium halobium and Haloferax volcanii*. *Mutat. Res.* **1996**, *364*, 25–32. [CrossRef]

9. McCready, S.; Marcello, L. Repair of UV damage in *Halobacterium salinarum*. *Biochem. Soc. Trans.* **2003**, *31*, 694–698. [CrossRef] [PubMed]

10. Crowley, D.J.; Boubriak, I.; Berquist, B.R.; Clark, M.; Richard, E.; Sullivan, L.; DasSarma, S.; McCready, S. The *uvrA, uvrB* and *uvrC* genes are required for repair of ultraviolet light induced DNA photoproducts in *Halobacterium sp.* NRC-1. *Saline Syst.* **2006**, *2*, 11. [CrossRef] [PubMed]

11. Stantial, N.; Dumpe, J.; Pietrosimone, K.; Baltazar, F.; Crowley, D.J. Transcription-coupled repair of UV damage in the halophilic archaea. *DNA Repair (Amst)* **2016**, *11*, 63–68. [CrossRef] [PubMed]

12. Sharma, N.; Hepburn, D.; Fitt, P.S. Photoreactivation in pigmented and non-pigmented extreme halophiles. *Biochim. Biophys. Acta* **1984**, *799*, 135–142. [CrossRef]

13. Shahmohammadi, H.R.; Asgarani, E.; Terato, H.; Saito, T.; Ohyama, Y.; Gekko, K.; Yamamoto, O.; Ide, H. Protective roles of bacterioruberin and intracellular KCl in the resistance of *Halobacterium salinarium* against DNA-Damaging agents. *J. Radiat. Res.* **1998**, *39*, 251–262. [CrossRef] [PubMed]

14. Kelly, M.; Norgard, S.; Liaaen-Jensen, S. C50-carotenoids. 5. carotenoids of *Halobacterium Salinarium*, especially bacterioruberin. *Acta Chem. Scand.* **1970**, *24*, 2169–2182. [CrossRef] [PubMed]

15. Kushwaha, S.C.; Gochnauer, M.B.; Kushner, D.J.; Kates, M. Pigments and isoprenoid compounds in extremely and moderately halophilic bacteria. *Can. J. Microbiol.* **1974**, *20*, 241–245. [CrossRef] [PubMed]

16. Kushwaha, S.C.; Kramer, J.K.G.; Kates, M. Isolation and characterization of C50-carotenoid pigments and other polar isoprenoids from *Halobacterium cutirubrum*. *Biochim. Biophys. Acta* **1975**, *398*, 303–314. [CrossRef]

17. Marshall, C.P.; Leuko, S.; Coyle, C.M.; Walter, M.R.; Burns, B.P.; Neilan, B.A. Carotenoid analysis of halophilic archaea by resonance raman spectroscopy. *Astrobiology* **2007**, *7*, 631–643. [CrossRef] [PubMed]

18. DasSarma, S.; Kennedy, S.P.; Berquist, B.; Ng, W.V.; Baliga, N.S.; Spudich, J.L.; Krebs, M.P.; Eisen, J.A.; Johnson, E.A.; Johnson, C.H.; Hood, L. Genomic perspective on the photobiology of *Halobacterium* species NRC-1, a phototrophic, phototactic, and UV-tolerant haloarchaeon. *Photosynth. Res.* **2001**, *70*, 3–17. [CrossRef] [PubMed]

19. Oesterhelt, D.; Stoekenius, W. Functions of a new photoreceptor membrane. *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 2853–2857. [CrossRef] [PubMed]

20. Papke, R.T.; Douady, C.J.; Doolittle, W.F.; Rodriguez-Valera, F. Diversity of bacteriorhodopsins in different hypersaline waters from a single Spanish saltern. *Environ. Microbiol.* **2003**, *5*, 1039–1045. [CrossRef] [PubMed]

21. NCBI Genome Database. Available online: http://www.ncbi.nlm.nih.gov/genome/browse/# (accessed on 23 August 2016).

22. Haynes, R.H. Molecular localization of radiation damage relevant to bacterial inactivation. In *Physical Processes in Radiation Biology*; Augenstein, L., Mason, R., Rosenberg, B., Eds.; Academic Press: New York, NY, USA, 1964; p. 51.

23. Setlow, R.B. Cyclobutane-type pyrimidine dimers in polynucleotides. *Science* **1966**, *153*, 379–386. [CrossRef] [PubMed]

24. Setlow, R.B.; Carrier, W.L. Pyrimidine dimers in ultraviolet-Irradiated DNA's. *J. Mol. Biol.* **1966**, *17*, 237–254. [CrossRef]

25. Joux, F.; Jeffrey, W.H.; Lebaron, P.; Mitchell, D.L. Marine bacterial isolates display diverse responses to UV-B radiation. *Appl. Environ. Microbiol.* **1999**, *65*, 93820–93827.

26. Kennedy, S.P.; Ng, W.V.; Salzberg, S.L.; Hood, L.; DasSarma, S. Understanding the adaptation of Halobacterium species NRC-1 to its extreme environment through computational analysis of its genome sequence. *Genome. Res.* **2001**, *11*, 1641–1650. [CrossRef] [PubMed]

27. Douki, T.; Cadet, J. Individual determination of the yield of the main UV-Induced dimeric pyrimidine photoproducts in DNA suggests a high mutagenicity of CC photolesions. *Biochemistry* **2001**, *40*, 2495–2501. [CrossRef] [PubMed]

28. Sinha, R.P.; Häder, D.P. UV-induced DNA damage and repair: A review. *Photochem. Photobiol. Sci.* **2002**, *1*, 225–236. [CrossRef] [PubMed]

29. Besaratinia, A.; Yoon, J.; Schroeder, C.; Bradforth, S.E.; Cockburn, M.; Pfeifer, G.P. Wavelength dependence of ultraviolet radiation-induced DNA damage as determined by laser irradiation suggests that cyclobutane pyrimidine dimers are the principal DNA lesions produced by terrestrial sunlight. *FASEB J.* **2011**, *25*, 3079–3091. [CrossRef] [PubMed]

30. Mitchell, D.L.; Jen, J.; Cleaver, J.E. Sequence specificity of cyclobutane pyrimidine dimers in DNA treated with solar (ultraviolet B) radiation. *Nucleic Acids Res.* **1991**, *20*, 225–229. [CrossRef]

31. Yoon, J.H.; Lee, C.S.; O'Connor, T.R.; Yasui, A.; Pfeifer, G.P. The DNA damage spectrum produced by simulated sunlight. *J. Mol. Biol.* **2000**, *299*, 681–693. [CrossRef] [PubMed]

32. Rastogi, R.P.; Richa; Kumar, A.; Tyagi, M.B.; Sinha, R.P. Molecular Mechanisms of Ultraviolet Radiation-Induced DNA Damage and Repair. *J. Nucleic Acids* **2010**, *2010*, 1–32. [CrossRef] [PubMed]

33. Matallana-Surget, S.; Meador, J.A.; Joux, F.; Douki, T. Effect of the GC content of DNA on the distribution of UVB-induced bipyrimidine photoproducts. *Photochem. Photobiol. Sci.* **2008**, *7*, 794–801. [CrossRef] [PubMed]

34. Jones, D.L. Dinucleotide Counts. Available online: https://github.com/jones1040/DinucleotideCounts (accessed on 23 August 2016).

35. R: A Language and Environment for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 23 August 2016).

36. Hothorn, T.; Bretz, F.; Westfall, P. Simultaneous inference in general parametric models. *Biometrical. J.* **2008**, *50*, 346–363. [CrossRef] [PubMed]

37. Bolhuis, H.; Palm, P.; Wende, A.; Falb, M.; Rampp, M.; Rodriguez-Valera, F.; Pfeiffer, F.; Oesterhelt, D. The genome of the square archaeon *Haloquadratum walsbyi*: Life at the limits of water activity. *BMC Genom.* **2006**, *7*, 169. [CrossRef] [PubMed]

38. Burns, D.G.; Janssen, P.H.; Itoh, T.; Kamekura, M.; Li, Z.; Jensen, G.; Rodrıguez-Valera, F.; Bolhuis, H.; Dyall-Smith, M. *Haloquadratum walsbyi* gen. nov., sp. nov., the square haloarchaeon of Walsby, isolated from saltern crystallizers in Australia and Spain. *Int. J. Syst. Evolut. Microbiol.* **2007**, *57*, 387–392. [CrossRef] [PubMed]

39. Porter, K.; Russ, B.; Burns, D.; Dyall-Smith, M. Diversity of *Haloquadratum* and other haloarchaea in three, geographically distant, Australian saltern crystallizer ponds. *Extremophiles* **2010**, *14*, 161–169.

40. Sancar, G.B. Enzymatic photoreactivation: 50 years and counting. *Mutat. Res.* **2000**, *451*, 25–37. [CrossRef]

41. Weber, S. Light-driven enzymatic catalysis of DNA repair: A review of recent biophysical studies on photolyase. *Biochim. Biophys. Acta* **2005**, *1707*, 1–23. [CrossRef] [PubMed]

42. Sancar, A. DNA excision repair. *Annu Rev. Biochem.* **1996**, *65*, 43–81. [CrossRef] [PubMed]

43. McKenzie, R.L.; Björn, L.O.; Bais, A.; Ilyasd, M. Changes in biologically active ultraviolet radiation reaching the Earth's surface. *Photochem. Photobiol. Sci.* **2003**, *2*, 5–15. [CrossRef] [PubMed]

44. Chou, M.D.; Lee, K.T. Parameterizations for the absorption of solar radiation by water vapor and ozone. *J. Atmos. Sci.* **1996**, *53*, 1203–1208. [CrossRef]

45. Litchfield, C.D. Survival strategies for microorganisms in hypersaline environments and their relevance to life on early Mars. *Meteorit. Planet Sci.* **1998**, *33*, 813–819. [CrossRef] [PubMed]

46. Paul, S.; Bag, S.K.; Das, S.; Harvill, E.T.; Dutta, C. Molecular signature of hypersaline adaptation: Insights from genome and proteome composition of halophilic prokaryotes. *Genome Biol.* **2008**, *9*, R70. [CrossRef] [PubMed]

47. Knight, R.D.; Freeland, S.J.; Landweber, L.F. A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. *Genome Biol.* **2001**, *2*, research0010.1–0010.13. [CrossRef]

48. Musto, H.; Naya, H.; Zavala, A.; Romero, H.; Alvarez-Valín, F.; Bernardi, G. Genomic GC level, optimal growth temperature, and genome size in prokaryotes. *Biochem. Biophys. Res. Commun.* **2006**, *347*, 1–3. [CrossRef] [PubMed]

49. Hildebrand, F.; Meyer, A.; Eyre-Walker, A. Evidence of selection upon genomic GC-content in bacteria. *PLoS Genet.* **2010**, *6*, e1001107. [CrossRef] [PubMed]