Article

In Silico Prophage Analysis of *Halobacterium salinarum* ATCC 33170

Danielle L. Peters [1],*, Bassel Akache [1], Wangxue Chen [1,2] and Michael J. McCluskie [1]

1 Human Health Therapeutics (HHT) Research Center, National Research Council Canada, Ottawa, ON K1A 0R6, Canada
2 Department of Biology, Brock University, St. Catharines, ON L2S 3A1, Canada
* Correspondence: danielle.peters@nrc-cnrc.gc.ca

Abstract: The extremophile *Halobacterium salinarum* is an aerobic archaeon that has adapted to thrive in high-salt environments such as salted fish, hypersaline lakes, and salterns. Halophiles have garnered significant interest due to their unique interactions with bacteriophages known as haloarchaeophages. Studies have identified and characterized prophages in halophilic archaea, such as *Haloferax volcanii*, *Haloquadratum walsbyi*, and *Haloarcula marismortui*. Still, an investigation has yet to be conducted into the presence of prophage elements on *Halobacterium salinarum* ATCC 33170. This is of particular interest to us as we are using this strain as a source of archaeol, as one of the components of our sulfated lactosyl archaeol (SLA) archaeosome adjuvant. Genomic contigs of strain 33170 were bioinformatically assessed for prophage-like features using BLAST, PHASTER, InterProScan, and PHYRE2. A 7 kb region encoding six genes was identified as an incomplete prophage, and the proteins were further analyzed, revealing high homology to proteins encoded by bacteria, archaea, and an IS200 transposon. Restricting the BLASTp database to viruses resulted in hits to both myo- and siphoviral proteins, which would be unusual for an intact prophage. Additionally, no known phage structural proteins were identified in the search, suggesting a low chance that *H. salinarum* ATCC 33170 harbors a latent prophage.

Keywords: archaea; prophage; halophile; bacteriophage; temperate phage; haloarchaeophages

1. Introduction

Halobacterium salinarum is an aerobic archaeon found in hypersaline environments, such as salted fish, salt flats, and salterns. These archaea play a crucial role in the ecology and biogeochemical cycles of hypersaline environments and, thus, have garnered significant attention due to the unique adaptations they have to thrive in high-salt conditions. These adaptations include the production of compatible solutes to maintain osmotic balance and modified membrane proteins [1]. Furthermore, *H. salinarum* can perform phototrophy using bacteriorhodopsin, a reddish light-driven proton pump. Due to this archaeon's extreme conditions, it possesses robust DNA repair mechanisms to counteract the damaging effects of high salt and UV radiation, ensuring genome stability in harsh environments [2]. We are also particularly interested in *H. salinarum* as a source of archaeol, one of the main components of our sulfated lactosyl archaeol (SLA) archaeosome adjuvant. Indeed, SLA archaeosomes are liposomal vesicles composed of a sulfated disaccharide group covalently linked to the free sn-1 hydroxyl backbone of an archaeal core lipid derived from *H. salinarum* [3]. This adjuvant has shown great promise in pre-clinical studies, and a better understanding of potential prophage elements within *H. salinarum* is necessary to support the continued progression of SLA archaeosomes toward clinical applications [4–6]. The prophage status of this strain is important, as culturing this organism in a good manufacturing practice (GMP) facility requires the absence of active prophages to eliminate phage contamination risks to the facility and resulting products.

Extremophiles such as *H. salinarum* thrive in harsh environments and maintain intricate interactions with viruses that target archaea. Haloarchaeophages are viruses that specifically infect halophilic archaea [7]. These viruses exhibit a wide range of morphologies, including tailed and non-tailed phages, with double-stranded DNA genomes being the most common [7]. They possess specialized mechanisms to attach to and inject their genetic material into host cells, leading to the replication and assembly of new virus particles. Haloarchaeophages have demonstrated high host specificity, selectively infecting particular species or strains of halophilic archaea.

The study of haloarchaeophages has provided valuable insights into the biology of halophiles and their viral interactions [7,8]. These viruses influence the population dynamics and diversity of halophiles, serving as agents of selection and contributing to the evolution of their hosts. Moreover, the infection and lysis of halophilic archaea by haloarchaeophages release organic matter and nutrients back into the environment, influencing the biogeochemical cycles of hypersaline habitats.

Prophages are viral genomes integrated into the host genome, remaining dormant until triggered to enter the lytic cycle. Studies have identified prophages in halophilic archaea, such as *Haloferax volcanii*, *Haloquadratum walsbyi*, and *Halobacterium halobium*, among others [9–11]. Prophage elements in halophile genomes indicate that viruses can integrate their genetic material into the host genome and coexist with their hosts for extended periods. Under certain conditions, such as stress or changes in the host cell environment, the prophage may be induced to enter the lytic cycle, producing new virus particles and lysis of the host cell. The study of prophages in halophiles provides insights into the genetic diversity and evolution of halophilic archaea. It also contributes to understanding viral–host interactions in extreme environments and the mechanisms underlying viral latency and reactivation.

Haloarchaeophages also hold promise in various biotechnological applications. Their stability in high salt concentrations and extreme conditions makes them valuable tools for genetic engineering, DNA delivery systems, and bioremediation. The unique properties of haloarchaeophages offer opportunities to develop innovative molecular biology and biotechnology approaches.

This study aims to perform a comprehensive genetic analysis of *Halobacterium salinarum* strain ATCC 33170 to identify any prophage sequences encoded within its genome. Advanced bioinformatic analyses have been employed to identify putative prophage sequences within the ATCC 33170 genome. Functional annotation and prediction of prophage-associated genes provide insights into potential roles in host–virus interactions, such as lysogeny, host manipulation, or environmental adaptation. The findings of this study contribute to our understanding of the genetic landscape of *Halobacterium salinarum* strains and their viral interactions. Unraveling the prophage sequences and their potential functions will provide valuable insights into the coevolutionary dynamics between halophilic archaea and viruses in highly saline environments and provide data to support the further development of components, such as archaeol, derived from halophiles.

2. Materials and Methods

Genomic data for *H. salinarum* ATCC 33170 were downloaded from ATCC and NCBI (NZ_JACHGX010000001—NZ_JACHGX010000037) for prophage analysis. The ATCC genomic data consist of 28 contiguous overlapping DNA segments (contigs), with the largest contig measuring 718.6 kilobases (kb). The NCBI genomic data consist of 37 contigs, with the largest contig measuring 642.1 kb long. The genomic data were imported into Geneious Prime and used for de novo assembly with Flye version 2.7 [12]. Input data type was selected as high-quality contigs with an estimated genome size of 2.4 megabases (Mb). No sequences were trimmed before assembly.

The Flye contigs were split into 50 and 90 kb sections and submitted to BLASTX and BLASTN (discontiguous megablast) using the non-redundant protein database (accessed January 2022). The first submission used the nr database restricted to viruses (tax id: 10239),

then analyzed again without database restrictions to identify regions with high viral hits spanning more than 5 kb.

The contig file stemming from the Flye assembly was also submitted to PHASTER (www.phaster.ca, accessed on 13 January 2022) with the checkbox for submitting a file consisting of multiple separate contigs selected [13]. The results were downloaded and indicated the presence of an incomplete prophage region on contig 14. An in-depth analysis of this region was conducted in Geneious Prime. The study revealed that the incomplete prophage region contains six genes translated to obtain protein sequences for further analysis with BLASTP, PHYRE2, and InterProScan [14–16].

The six proteins were submitted to BLASTP against the non-redundant protein database and a restricted viral database, and the top hit was documented. The proteins were also submitted to PHYRE2 to identify the closest protein templates based on the predicted structure [14]. Results were downloaded and analyzed for protein functionality and origin.

The InterProScan [15,16] plugin for Geneious Prime was used to analyze the protein sequences against the following member databases: Conserved Domains Database (CDD [17]), Gene3d [18], High-quality Automated and Manual Annotation of Proteins (HAMAP [19]), Panther [20], PfamA [21], Protein Information Resource SuperFamily (PIRSF [22]), Simple Modular Architecture Research Tool (SMART [23]), SuperFamily [24], and The Institute for Genomic Research collection of manually curated protein families (TIGRFAM [25]). Results were annotated to the nucleotide sequence of the incomplete prophage region, and protein families were analyzed for taxonomic associations.

The contig nucleotide sequences were submitted to Prokaryotic Antiviral Defence LOCator (PADLOC) v2.0.0 to identify any potential antiviral defense systems encoded [26].

3. Results

3.1. Scaffold Construction

Flye Contig Assembly

Flye assembled the collective 65 ATCC and NCBI contigs into 21 contigs with a length sum of 2.4 Mb (Table 1). The longest contig was 1.03 Mb, and the assembly had an N50 length of 718,706 base pairs (bp).

Table 1. Flye assembly statistics.

Statistics	All Contigs
Number of contigs	21
Min Length (bp)	1323
Median Length (bp)	12,424
Mean Length (bp)	115,285
Max Length (bp)	1,030,011
N50 Length (bp)	718,706

3.2. Genomic Investigations of H. salinarum ATCC 33170 Contigs for Prophage-like Elements

3.2.1. PHASTER

PHASTER analysis identified a single incomplete prophage region, 7 kb in length from bases 3957–11,025 on contig 14, encoding six proteins (Table 2). The prophage completeness score given was 40, with the scoring scale described as follows: incomplete prophage cut-off < 70, questionable prophage cut-off 70–90, and intact prophage cut-off > 90. The resulting PHASTER BLAST hits of this region are listed in Table 2 and include two hypothetical proteins of unknown function (gene product (gp)) 1 and gp3, a putative pore-forming tail tip protein (gp2), two putative transposases (gp4 and gp5), and insertion sequence element Dka2 orfB (gp6).

Table 2. PHASTER results from incomplete prophage region of contig 14.

Gp	CDS Position	BLAST Hit	E-Value
1	3957–4709 *	PHAGE_Sphing_PAU_NC_019521: gp187; PP_02344; phage(gi435844690)	5.71×10^{-27}
2	4773–6065	PHAGE_Vibrio_vB_VchM_Kuja_NC_048827: pore-forming tail tip protein; PP_02345; phage(gi100139)	1.02×10^{-40}
3	6244–7476 *	PHAGE_Archae_virus_NC_008695: hypothetical protein; PP_02346; phage(gi119757001)	7.73×10^{-15}
4	7705–8961 *	PHAGE_Rhodot_RM378_NC_004735: putative transposase; PP_02347; phage(gi30044048)	3.25×10^{-16}
5	8963–9355 *	PHAGE_Sulfol_SMV3_NC_029103: putative transposase IS200-family protein; PP_02348; phage(gi985760792)	1.14×10^{-32}
6	9745–11025 *	PHAGE_Hypert_2_NC_014321: IS element Dka2 orfB; PP_02349; phage(gi300116745)	2.06×10^{-17}

* Reverse complement.

Examination of the PHASTER hits for taxonomic association reveals that each protein belongs to a different phage family (Table 3). The taxonomic hits are *Myoviridae* (2), *Ackermannviridae*, *Siphoviridae*, *Bicaudaviridae*, and an unclassified archaeal virus. The phage genome lengths for these phages ranged from 17,666 bp to 219,372 bp, significantly larger than the 7 kb region identified as an incomplete prophage.

Table 3. Taxonomic information on PHASTER hits for each protein.

Gp	Phage	Accession	Taxonomy	Genome Length (bp)
1	*Sphingomonas* phage PAU	NC_019521	*Myoviridae*	219,372
2	Vibrio phage vB_VchM_Kuja	NC_048827.1	*Ackermannviridae*	148,180
3	Archaeal BJ1 virus	NC_008695.1	*Siphoviridae*	42,271
4	*Rhodothermus* phage RM378	NC_004735.1	*Myoviridae*	129,908
5	*Sulfolobus monocaudavirus* SMV3	NC_029103	*Bicaudaviridae*	64,323
6	Hyperthermophilic Archaeal Virus 2	NC_014321	Unclassified archaeal virus	17,666

Further investigation into gp2, which was initially identified as matching the putative pore-forming tail tip protein of Vibrio phage vB_VchM_Kuja (Kuja), was completed using BLASTP against virally-restricted and un-restricted databases. The resulting protein hits are a UDP-glucose 6-dehydrogenase (Tables 4 and 5). As such, it appears that the Kuja protein is incorrectly labeled in the PHASTER database, as it is labeled as UDP-glucose 6-dehydrogenase in NCBI and encoded in the metabolism region of the genome.

Table 4. BLASTP results from a virus-restricted, non-redundant database search.

Gp	Description	Scientific Name	Morphology	Coverage (%)	E-Value	ID (%)	Accession Number
1	gp187	*Sphingomonas* phage PAU	Myovirus	79	2.00×10^{-25}	32.67	YP_007006794.1
2	TPA: MAG TPA: putative UDP-glucose 6-dehydrogenase	*Siphoviridae* sp. ctCCX1	Siphovirus	98	1.00×10^{-91}	38.27	DAD67937.1
3	Orc1-type DNA replication protein	*Halobacterium* virus ChaoS9	Myovirus	89	5.00×10^{-11}	25.13	YP_010078015.1
4	Transposase	*Halobacterium* phage phiH	Myovirus	100	0	100.00	YP_009981874.1
5	IS200/IS605-like element ISH1-8 family transposase	*Halobacterium* phage phiH	Myovirus	100	2.00×10^{-91}	100.00	YP_009981873.1
6	TPA: MAG TPA: putative transposase	*Siphoviridae* sp.	Siphovirus	84	8.00×10^{-26}	25.65	DAT22193.1

Table 5. Top PHYRE2 hit per gene product from the incomplete prophage region of contig 14.

Gp	PDB Header; ID	PDB Molecule	PDB Title	Confidence (%)	Coverage (%)
1	Transferase; 2PA4	UTP-glucose-1-phosphate uridylyltransferase	Crystal structure of UDP-glucose pyrophosphorylase from *Corynebacteria glutamicum* in complex with magnesium and UDP-glucose	100	95
2	Oxidoreductase; 3VTF	UDP-glucose 6-dehydrogenase	Structure of UDP-glucose dehydrogenase from the hyperthermophilic archaeon *Pyrobaculum islandicum*	100	97
3	Replication/DNA; 2QBY	Cell division control protein 6 homolog 1	Crystal structure of a heterodimer of Cdc6/Orc1 initiators bound to origin DNA from *Sulfolobus solfataricus*	100	89
4	RNA binding protein/RNA/DNA; 7C7L	CRISPR-associated protein Cas14a.1	Cryo-EM structure of the Cas12f1-sgRNA-target DNA complex	100	82
5	Ferredoxin-like fold; 2F5G	Transposase IS200-like superfamily	Transposase IS200-like family	100	95
6	RNA binding protein/RNA/DNA; 7C7L	CRISPR-associated protein Cas14a.1	Cryo-EM structure of the Cas12f1-sgRNA-target DNA complex	100	85

3.2.2. BLAST

Analysis of all contigs with BLASTX and discontiguous megablast on an unrestricted database produced hits exclusive to halophiles without phage or prophage hits present. With the database restricted to viruses, hits are randomly spaced throughout the genome

without long stretches (>4 kb) of sequential phage hits. Focus on the incomplete prophage region of contig 14 using BLASTP with the unrestricted database shows 100% identity and 94–100% coverage against *Halobacterium* spp. (Supplementary Table S1). When the database is restricted to viruses, the gp hits within the incomplete prophage region are myoviral and siphoviral phages (Table 4). Two notable hits are for gp4 and gp5, with 100% coverage and identity to the *Halobacterium* phage phiH. These two phiH proteins are predicted to encode a transposase.

3.2.3. PHYRE2

The six protein sequences from the predicted incomplete prophage region of contig 14 were submitted to PHYRE2 to model the proteins and compare them to known structural templates. The top Protein Data Bank (PDB) hits for gp1 through gp3 are for bacterial and archaeon proteins with 100% confidence (Table 5). The top PDB hit for gp1 is a UTP-glucose-1-phosphate uridylyltransferase from the bacterium *Corynebacteria glutamicum* with 95% coverage. The enzyme UDP-glucose 6- dehydrogenase from the hyperthermophilic archaebacteria *Pyrobaculum islandicum* is the top PDB hit for gp2 with 97% coverage. Gp3 has 89% coverage to the cell division control protein 6 homolog 1 (Cdc6/Orc1) from the thermophilic archaeon *Sulfolobus solfataricus*.

The top PDB hits for gp4 and gp6 were to an uncultured archaeon Cas14a.1 protein with 82 and 85% coverage, respectively, and 100% confidence. These gene products also had a single hit to bacteriophage T7 DNA primase/helicase. The gp4 T7 hit ranked 93 of 120, with 47.9% confidence and 16% coverage. The gp6 T7 hit ranked 80 of 120, with 82% confidence and 20% coverage. Finally, the gp5 top hit was a transposase from the IS200-like superfamily of insertion sequences with 100% confidence and 95% coverage. Hit 30 of 69 in the gp5 results was to a phage replication organizer domain, with 8.9% confidence and 25% coverage.

3.2.4. InterProScan

Significant database hits were obtained with InterProScan against the six proteins of interest (Table 6). The database hits confirm the results obtained with PHYRE2 and BLASTP.

Table 6. InterProScan hit results per gene product from the incomplete prophage region of contig 14.

Gp	Name	Type	Id	InterPro ID
	NTP_transferase	CDD	cd04181	
	Spore coat polysaccharide biosynthesis protein SpsA; chain A	Gene3D	G3DSA:3.90.550.10	IPR029044
1	N-acetylmuramate alpha-1-phosphate uridylyltransferase	Panther	PTHR43584:SF6	
	NTP_transferase	Pfam	PF00483	IPR005835
	Nucleotide-diphospho-sugar transferases	Superfamily	SSF53448	IPR029044

Table 6. *Cont.*

Gp	Name	Type	Id	InterPro ID
2	NAD(P)-binding Rossmann-like domain	Gene3D	G3DSA:3.40.50.720	
	Cytochrome C1, transmembrane anchor, C-terminal	Gene3D	G3DSA:1.20.5.100	
	UDP-glucose 6-dehydrogenase tuad	Panther	PTHR43750	
	UDPG_MGDP_dh_N	Pfam	PF03721	IPR001732
	UDPG_MGDP_dh_C	Pfam	PF03720	IPR014027
	UDPG_MGDP_dh	Pfam	PF00984	IPR014026
	UDPglc_GDPman_dh	PIRSF	PIRSF000124	IPR017476
	UDPglc_DH_bac	PIRSF	PIRSF500134	IPR028357
	UDPG_MGDP_dh_C_a_2_a	SMART	SM00984	IPR014027
	NAD(P)-binding Rossmann-fold domains	Superfamily	SSF51735	IPR036291
	UDP-glucose/GDP-mannose dehydrogenase C-terminal domain	Superfamily	SSF52413	IPR036220
	6-phosphogluconate dehydrogenase C-terminal domain-like	Superfamily	SSF48179	IPR008927
	TIGR03026	TIGRFAM	TIGR03026	IPR017476
3	AAA	CDD	cd00009	
	Cdc6_C	CDD	cd08768	IPR015163
	Helicase, RuvA protein; domain 3	Gene3D	G3DSA:1.10.8.60	
	P-loop containing nucleoside triphosphate hydrolases	Gene3D	G3DSA:3.40.50.300	IPR027417
	Winged helix-like DNA-binding domain superfamily	Gene3D	G3DSA:1.10.10.10	IPR036388
	Orc1_type_DNA_replic_protein	HAMAP	MF_01407	IPR014277
	Orc1-type DNA replication protein 1	Panther	PTHR10763:SF22	
	AAA_22	Pfam	PF13401	IPR003593
	Cdc6_C	Pfam	PF09079	IPR015163
	AAA_5	SMART	SM00382	IPR003593
	Cdc6_C_2	SMART	SM01074	IPR015163
	P-loop containing nucleoside triphosphate hydrolases	Superfamily	SSF52540	IPR027417
	Winged helix DNA-binding domain	Superfamily	SSF46785	IPR036390
	TIGR02928	TIGRFAM	TIGR02928	IPR014277
4	Transposase	Panther	PTHR30405	
	Neutral protease	Panther	PTHR30405:SF17	
	OrfB_Zn_ribbon	Pfam	PF07282	IPR010095
	Transposase, IS605 OrfB family, central region	TIGRFAM	TIGR01766	IPR010095

Table 6. *Cont.*

Gp	Name	Type	Id	InterPro ID
5	Transposase IS200-like superfamily	Gene3D	G3DSA:3.30.70.1290	IPR036515
	IS200-like transposase	Panther	PTHR33360:SF3	
	Y1_Tnp	Pfam	PF01797	IPR002686
	Y1_Tnp_2	SMART	SM01321	IPR002686
	Transposase IS200-like	Superfamily	SSF143422	IPR036515
6	Neutral protease	Panther	PTHR30405:SF17	
	OrfB_IS605	Pfam	PF01385	IPR001959
	OrfB_Zn_ribbon	Pfam	PF07282	IPR010095
	Transposase, IS605 OrfB family, central region	TIGRFAM	TIGR01766	IPR010095

3.2.5. PADLOC

The contigs were analyzed with PADLOC to identify any putative antiviral defense systems encoded. There were six hits across five contigs (Table 7). Contig 1 was found to encode PDC-S70, which is a putative Phage Defense Candidate with an unknown function [26]. Contig 8 encodes the HEC-05 system, a helicase, methylase, ATPase (Hma)-embedded candidate [26]. There are two SoFic proteins encoded on contigs 16 and 17, which ligates AMP onto target proteins [27,28]. Finally, contig 19 encodes two putative DNA-modification system proteins (DMS_other) labeled Specificity_I and REase_I (Table 7). The PADLOC database warns DMS_other hits, comprised of various DNA-modification system proteins, such as restriction-modification, Bacteriophage Exclusion (BREX) [29], and Defense Island System Associated with Restriction-Modification (DISARM) [30], is generated with a permissive model that may not return genuine phage defense systems and should therefore be treated with caution [31].

Table 7. Putative antiviral defense systems encoded by *H. salinarum* identified using PADLOC.

Contig	System	Hidden Markov Model Accession	Hidden Markov Model Name	Domain E Value	Coverage	Start	End	Strand *
1	PDC-S70	PDLC05071	PDC-S70_WP_028295898.1	1.20×10^{-9}	0.969	4241	4753	-
8	HEC-05	PDLC04151	HEC-05_WP_156276309.1	2.20×10^{-166}	0.953	197190	199046	-
16	SoFic	PDLC03963	SoFic__SoFic	1.40×10^{-94}	0.972	31762	33003	+
17	SoFic	PDLC03963	SoFic__SoFic	2.00×10^{-95}	0.972	17786	19027	-
19	DMS_other	PDLC03108	Specificity_I_00057	2.00×10^{-214}	0.998	973	2400	+
19	DMS_other	PDLC03040	REase_I_00001	2.70×10^{-118}	0.977	2479	5454	+

* Strand refers to the open reading frame being encoded on the forward strand (+) or reverse strand (-).

4. Discussion

Halophilic euryarchaeotes, a group of extremophilic archaea adapted to high-salt environments, have attracted considerable scientific interest due to their ecological significance and unique adaptations. They play a crucial role in the biogeochemical cycling of hypersaline environments and offer valuable insights into the limits of life on Earth [32–34]. Hypersaline waters and salt crystals contain high numbers of haloarchaeal cells and their viruses, representing a worldwide distributed reservoir of orphan genes and possibly novel virion morphotypes [7,8]. The study of haloarchaeal-associated viruses, known as halophages, provides a deeper understanding of viral–host interactions and unveils potential biotechnological applications due to their unique features [8].

Over 110 viruses have been described for halophilic archaeal hosts with lifestyles ranging from lytic to forming chronic infections of their hosts [7,9,35]. Prophages or prophage-like elements have been documented in halophilic archaeal genomes, such as *Haloferax volcanii*, *Haloquadratum walsbyi*, and *Halobacterium halobium* [9–11]. Most of these prophage sequences correspond to pleomorphic viruses, which comprise the second largest group of halophages, with over 18 documented [9,35]. These viruses establish non-lytic infections, presumably using budding as an exit mechanism, and have been isolated against archaea genera, such as *Halorubrum*, *Haloarcula*, *Halogeometricum*, and *Natrinema* [9,35–39].

One seemingly pervasive type of pleomorphic halophage is the *Haloarcula* His2, which shares protein similarity with several prophages in the halophiles [40–42]. For example, *Haloquadratum walsbyi* was found to encode an incomplete prophage related to His2 and two pleomorphic haloviruses, HRPV-1 and HHPV1, which also have a similar block of homologs related to His2 [10,35,42]. Furthermore, the *Halorubrum* pleomorphic virus-1 (HRPV-1), isolated from a solar saltern, was found to encode three structural proteins, VP3, VP4, and VP8 [36,43]. The HRPV-1 encoded proteins show significant similarity to the proteins of the minimal replicon of plasmid pHK2 of *Haloferax* sp. and the His2 phage [36,42,44]. As several halophage sequences are available from a broad array of halophiles, analysis of the *H. salinarum* genome for the presence of prophage sequences is possible.

PHASTER analysis of the 21 Flye contigs identified a 7 kb portion of contig 14 encoding putative prophage elements. The assigned prophage completeness score was low (i.e., 40), suggesting an incomplete prophage region. Further analysis of all contigs using BLASTX and discontiguous megablast did not reveal other areas encoding potential prophage elements in the genome. While some of the genes appear to be viral in origin, BLASTP analysis with a restricted database shows mixed viral-family hits, which is highly unusual for an intact phage.

PHYRE2 analysis on the six proteins of interest revealed high homology to various proteins from bacteria, archaea, and a transposon (IS200) with 100% confidence, as discussed in the Results section. The gp1 protein was modeled to a glucose-1-phosphate uridylyltransferase, or UGPase, an enzyme that catalyzes UDP-glucose production from glucose-1-phosphate and UTP [45]. This enzyme is widespread due to its role in glycogen synthesis and forming glycolipids, glycoproteins, and proteoglycans [46–48]. Although glycoproteins are featured in some halophage capsids, this enzyme was not modeled to any phage proteins [43]. Gp2 PHYRE2 results are to UDP-glucose dehydrogenase, which catalyzes a two-step NAD-dependent oxidation of UDP-glucose (UDP-Glc) to produce UDP-glucuronic acid (UDP-GlcA) [49]. Studies of this enzyme have demonstrated its importance in polysaccharide biosynthesis and detoxification [49]. This enzyme was also not modeled to a protein with PHYRE2.

Modeling of the gp3 protein hit the crystal structure of a heterodimer of Cdc6/Orc1 initiators bound to the origin DNA from *Sulfolobus solfataricus* [50]. Cellular initiators form higher-order assemblies on replication origins, using ATP to remodel duplex DNA and facilitate the loading of replisome components [51]. This protein appears to be a core component of the basal initiation machinery used to recognize the origin of replication in *S. solfataricus*, suggesting it is not a phage-derived protein.

Further investigation of gp4 and gp6 showed that their structures are similar to the cryo-EM structure of the Cas12f1-sgRNA-target DNA complex [52]. The type V-F Cas12f proteins are compact and associate with a guide RNA to cleave single- and double-stranded DNA targets [53]. A cryo-electron microscopy structure revealed that two Cas12f1 molecules assemble with the single guide RNA to recognize the double-stranded DNA target [52,53]. Each Cas12f1 protomer plays distinct roles in nucleic acid recognition and DNA cleavage, explaining how the miniature enzyme achieves RNA-guided DNA cleavage [54,55]. There is a single hit to the T7 bacteriophage primase-2 helicase, though the modeled region is tiny in comparison and is ranked 93 and 80 for gp4 and gp6, respectively.

The gp5 protein is modeled to the crystal structure of IS200 transposase of *Sulfolobus solfataricus* [56]. IS200 transposases, present in many bacteria and Archaea, are distinct from other groups of transposases. Two monomers form a tight dimer, forming the catalytic site at the interface between the two monomers [56]. A phage hit corresponding to a DNA replication organizer membrane protein of phage Phi29 was identified [57]. The hit was ranked 30, having 9% confidence and 25% identity over 12 amino acids; thus, it is unlikely that gp5 is a phage-derived protein.

Examination of the contigs for phage-defense systems with PADLOC revealed six putative genes on five contigs. To defend against viruses, Archaea are thought to primarily use the Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated genes (CRISPR-Cas) system [58,59], Toxin–Antitoxin (TA) systems, Restriction Modification (RM) systems, and alteration of cell surface proteins [60,61]. No CRISPR-Cas or TA systems were identified in the analysis, only RM, AMP-ligating, and unidentified systems. Overall, the phage-defense system carriage rate of *H. salinarum* is low compared to bacteria, which can carry upwards of 15 defense systems [62].

As this strain will be used to produce adjuvants, the mutational behavior of nearby strains should be assessed. In the investigation of the evolutionary dynamics of *H. salinarum*, a mutation-accumulation experiment was conducted, and the genome was sequenced, comparing it to the moderate halophilic archaeon *Haloferax volcanii* [63]. The mutation accumulation in *H. salinarum* over 1250 generations revealed a base-substitution rate of 3.99×10^{-10} per site per generation, comparable to that of *H. volcanii* [63]. However, dissimilarities in genome-wide insertion–deletion rates and mutation spectra suggest unique evolutionary pathways. Notably, *H. salinarum* is characterized by a high rate of spontaneous mutations attributed to mobile genetic elements (MGEs), including ISH elements and transposons [64,65]. Studied since the 1980s, these elements are associated with the insertional inactivation of genes, as well as genome inversions and rearrangements. For example, differences between laboratory strains NRC-1 and R1 primarily stem from the dynamic mobilome, exemplifying the impact of MGEs on the evolutionary landscape of *H. salinarum* [65].

5. Conclusions

In conclusion, no genes encoding major phage structural proteins, such as capsid or coat proteins, were identified in the genome of *H. salinarum* ATCC 33170. An incomplete prophage region was identified using PHASTER, which was further investigated using BlastP, InterProScan, and PHYRE2. This suggests a lack of capacity to produce infectious phage particles in this archaeal strain. Additionally, proteins from the incomplete prophage region did not hit any viral proteins with significant confidence or coverage. The whole genome analysis of *Halobacterium salinarum* ATCC 33170 with various bioinformatics software programs suggests no intact prophage regions capable of producing functional virions are encoded. These findings support the safe production of *H. salinarum* ATCC 33170 in a GMP facility due to its low risk of active prophages, facilitating its use for commercial/medical applications. Further, studying halophiles and their interactions with haloarchaeophages provides a fascinating insight into the adaptations and dynamics of life in highly saline environments. Further research in this field will enhance our understanding of these unique organisms and their viruses while uncovering their potential applications in diverse areas of science and technology.

References

1. Eichler, J. *Halobacterium salinarum*: Life with More than a Grain of Salt: This Article Is Part of the Microbe Profiles Collection. *Microbiology* **2023**, *169*, 001327. [CrossRef]
2. McCready, S.; Marcello, L. Repair of UV Damage in *Halobacterium salinarum*. *Biochem. Soc. Trans.* **2003**, *31*, 694–698. [CrossRef]
3. McCluskie, M.J.; Deschatelets, L.; Krishnan, L. Sulfated Archaeal Glycolipid Archaeosomes as a Safe and Effective Vaccine Adjuvant for Induction of Cell-Mediated Immunity. *Hum. Vaccines Immunother.* **2017**, *13*, 2772–2779. [CrossRef] [PubMed]
4. Akache, B.; Deschatelets, L.; Harrison, B.A.; Dudani, R.; Stark, F.C.; Jia, Y.; Landi, A.; Law, J.L.M.; Logan, M.; Hockman, D.; et al. Effect of Different Adjuvants on the Longevity and Strength of Humoral and Cellular Immune Responses to the HCV Envelope Glycoproteins. *Vaccines* **2019**, *7*, 204. [CrossRef] [PubMed]
5. Akache, B.; Renner, T.M.; Tran, A.; Deschatelets, L.; Dudani, R.; Harrison, B.A.; Duque, D.; Haukenfrers, J.; Rossotti, M.A.; Gaudreault, F.; et al. Immunogenic and Efficacious SARS-CoV-2 Vaccine Based on Resistin-Trimerized Spike Antigen SmT1 and SLA Archaeosome Adjuvant. *Sci. Rep.* **2021**, *11*, 21849. [CrossRef] [PubMed]
6. Renner, T.M.; Akache, B.; Stuible, M.; Rohani, N.; Cepero-Donates, Y.; Deschatelets, L.; Dudani, R.; Harrison, B.A.; Baardsnes, J.; Koyuturk, I.; et al. Tuning the Immune Response: Sulfated Archaeal Glycolipid Archaeosomes as an Effective Vaccine Adjuvant for Induction of Humoral and Cell-Mediated Immunity towards the SARS-CoV-2 Omicron Variant of Concern. *Front. Immunol.* **2023**, *14*, 1182556. [CrossRef]
7. Luk, A.W.S.; Williams, T.J.; Erdmann, S.; Papke, R.T.; Cavicchioli, R. Viruses of Haloarchaea. *Life* **2014**, *4*, 681–715. [CrossRef]
8. Atanasova, N.S.; Roine, E.; Oren, A.; Bamford, D.H.; Oksanen, H.M. Global Network of Specific Virus-Host Interactions in Hypersaline Environments: Haloviruses Are Promiscuous. *Environ. Microbiol.* **2012**, *14*, 426–440. [CrossRef]
9. Alarcón-Schumacher, T.; Naor, A.; Gophna, U.; Erdmann, S. Isolation of a Virus Causing a Chronic Infection in the Archaeal Model Organism *Haloferax volcanii* Reveals Antiviral Activities of a Provirus. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2205037119. [CrossRef]
10. Dyall-Smith, M.L.; Pfeiffer, F.; Klee, K.; Palm, P.; Gross, K.; Schuster, S.C.; Rampp, M.; Oesterhelt, D. Haloquadratum Walsbyi: Limited Diversity in a Global Pond. *PLoS ONE* **2011**, *6*, e20968. [CrossRef]
11. Ken, R.; Hackett, N.R. Halobacterium Halobium Strains Lysogenic for Phage Phi H Contain a Protein Resembling Coliphage Repressors. *J. Bacteriol.* **1991**, *173*, 955–960. [CrossRef] [PubMed]
12. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of Long, Error-Prone Reads Using Repeat Graphs. *Nat. Biotechnol.* **2019**, *37*, 540–546. [CrossRef]
13. Arndt, D.; Grant, J.R.; Marcu, A.; Sajed, T.; Pon, A.; Liang, Y.; Wishart, D.S. PHASTER: A Better, Faster Version of the PHAST Phage Search Tool. *Nucleic Acids Res.* **2016**, *44*, W16–W21. [CrossRef] [PubMed]
14. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [CrossRef] [PubMed]
15. Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-Scale Protein Function Classification. *Bioinformatics* **2014**, *30*, 1236–1240. [CrossRef] [PubMed]
16. Blum, M.; Chang, H.-Y.; Chuguransky, S.; Grego, T.; Kandasaamy, S.; Mitchell, A.; Nuka, G.; Paysan-Lafosse, T.; Qureshi, M.; Raj, S.; et al. The InterPro Protein Families and Domains Database: 20 Years On. *Nucleic Acids Res.* **2021**, *49*, D344–D354. [CrossRef] [PubMed]
17. Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Gonzales, N.R.; Gwadz, M.; Lu, S.; Marchler, G.H.; Song, J.S.; Thanki, N.; Yamashita, R.A.; et al. The Conserved Domain Database in 2023. *Nucleic Acids Res.* **2023**, *51*, D384–D388. [CrossRef] [PubMed]
18. Lees, J.; Yeats, C.; Perkins, J.; Sillitoe, I.; Rentzsch, R.; Dessailly, B.H.; Orengo, C. Gene3D: A Domain-Based Resource for Comparative Genomics, Functional Annotation and Protein Network Analysis. *Nucleic Acids Res.* **2012**, *40*, D465–D471. [CrossRef] [PubMed]
19. Pedruzzi, I.; Rivoire, C.; Auchincloss, A.H.; Coudert, E.; Keller, G.; de Castro, E.; Baratin, D.; Cuche, B.A.; Bougueleret, L.; Poux, S.; et al. HAMAP in 2015: Updates to the Protein Family Classification and Annotation System. *Nucleic Acids Res.* **2015**, *43*, D1064–D1070. [CrossRef]

20. Mi, H.; Ebert, D.; Muruganujan, A.; Mills, C.; Albou, L.-P.; Mushayamaha, T.; Thomas, P.D. PANTHER Version 16: A Revised Family Classification, Tree-Based Classification Tool, Enhancer Regions and Extensive API. *Nucleic Acids Res.* **2021**, *49*, D394–D403. [CrossRef]

21. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.L.; Tosatto, S.C.E.; Paladin, L.; Raj, S.; Richardson, L.J.; et al. Pfam: The Protein Families Database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419. [CrossRef] [PubMed]

22. Wu, C.H. PIRSF: Family Classification System at the Protein Information Resource. *Nucleic Acids Res.* **2004**, *32*, D112–D114. [CrossRef]

23. Letunic, I.; Khedkar, S.; Bork, P. SMART: Recent Updates, New Developments and Status in 2020. *Nucleic Acids Res.* **2021**, *49*, D458–D460. [CrossRef]

24. Gough, J.; Karplus, K.; Hughey, R.; Chothia, C. Assignment of Homology to Genome Sequences Using a Library of Hidden Markov Models That Represent All Proteins of Known Structure. *J. Mol. Biol.* **2001**, *313*, 903–919. [CrossRef]

25. Haft, D.H.; Loftus, B.J.; Richardson, D.L.; Yang, F.; Eisen, J.A.; Paulsen, I.T.; White, O. TIGRFAMs: A Protein Family Resource for the Functional Identification of Proteins. *Nucleic Acids Res.* **2001**, *29*, 41–43. [CrossRef] [PubMed]

26. Payne, L.J.; Meaden, S.; Mestre, M.R.; Palmer, C.; Toro, N.; Fineran, P.C.; Jackson, S.A. PADLOC: A Web Server for the Identification of Antiviral Defence Systems in Microbial Genomes. *Nucleic Acids Res.* **2022**, *50*, W541–W550. [CrossRef] [PubMed]

27. Harms, A.; Stanger, F.V.; Dehio, C. Biological Diversity and Molecular Plasticity of FIC Domain Proteins. *Annu. Rev. Microbiol.* **2016**, *70*, 341–360. [CrossRef]

28. Engel, P.; Goepfert, A.; Stanger, F.V.; Harms, A.; Schmidt, A.; Schirmer, T.; Dehio, C. Adenylylation Control by Intra- or Intermolecular Active-Site Obstruction in Fic Proteins. *Nature* **2012**, *482*, 107–110. [CrossRef] [PubMed]

29. Goldfarb, T.; Sberro, H.; Weinstock, E.; Cohen, O.; Doron, S.; Charpak-Amikam, Y.; Afik, S.; Ofir, G.; Sorek, R. BREX Is a Novel Phage Resistance System Widespread in Microbial Genomes. *EMBO J.* **2015**, *34*, 169–183. [CrossRef]

30. Bravo, J.P.K.; Aparicio-Maldonado, C.; Nobrega, F.L.; Brouns, S.J.J.; Taylor, D.W. Structural Basis for Broad Anti-Phage Immunity by DISARM. *Nat. Commun.* **2022**, *13*, 2987. [CrossRef]

31. Millman, A.; Melamed, S.; Leavitt, A.; Doron, S.; Bernheim, A.; Hör, J.; Garb, J.; Bechon, N.; Brandis, A.; Lopatina, A.; et al. An Expanded Arsenal of Immune Systems That Protect Bacteria from Phages. *Cell Host Microbe* **2022**, *30*, 1556–1569.e5. [CrossRef] [PubMed]

32. Jung, J.; Kim, J.-S.; Taffner, J.; Berg, G.; Ryu, C.-M. Archaea, Tiny Helpers of Land Plants. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 2494–2500. [CrossRef] [PubMed]

33. Torregrosa-Crespo, J.; Martínez-Espinosa, R.M.; Esclapez, J.; Bautista, V.; Pire, C.; Camacho, M.; Richardson, D.J.; Bonete, M.J. Anaerobic Metabolism in Haloferax Genus. In *Advances in Microbial Physiology*; Elsevier: Amsterdam, The Netherlands, 2016; Volume 68, pp. 41–85. ISBN 978-0-12-804823-8.

34. Hochstein, L.I.; Lang, F. Purification and Properties of a Dissimilatory Nitrate Reductase from Haloferax Denitrificans. *Arch. Biochem. Biophys.* **1991**, *288*, 380–385. [CrossRef]

35. Demina, T.A.; Oksanen, H.M. Pleomorphic Archaeal Viruses: The Family Pleolipoviridae Is Expanding by Seven New Species. *Arch. Virol.* **2020**, *165*, 2723–2731. [CrossRef]

36. Pietilä, M.K.; Roine, E.; Paulin, L.; Kalkkinen, N.; Bamford, D.H. An ssDNA Virus Infecting Archaea: A New Lineage of Viruses with a Membrane Envelope. *Mol. Microbiol.* **2009**, *72*, 307–319. [CrossRef]

37. Zhang, Z.; Liu, Y.; Wang, S.; Yang, D.; Cheng, Y.; Hu, J.; Chen, J.; Mei, Y.; Shen, P.; Bamford, D.H.; et al. Temperate Membrane-Containing Halophilic Archaeal Virus SNJ1 Has a Circular dsDNA Genome Identical to That of Plasmid pHH205. *Virology* **2012**, *434*, 233–241. [CrossRef] [PubMed]

38. Chen, B.; Chen, Z.; Wang, Y.; Gong, H.; Sima, L.; Wang, J.; Ouyang, S.; Gan, W.; Krupovic, M.; Chen, X.; et al. ORF4 of the Temperate Archaeal Virus SNJ1 Governs the Lysis-Lysogeny Switch and Superinfection Immunity. *J. Virol.* **2020**, *94*, e00841-20. [CrossRef]

39. Witte, A.; Baranyi, U.; Klein, R.; Sulzner, M.; Luo, C.; Wanner, G.; Krüger, D.H.; Lubitz, W. Characterization of *Natronobacterium magadii* Phage ΦCh1, a Unique Archaeal Phage Containing DNA and RNA. *Mol. Microbiol.* **1997**, *23*, 603–616. [CrossRef] [PubMed]

40. Demina, T.A.; Atanasova, N.S.; Pietilä, M.K.; Oksanen, H.M.; Bamford, D.H. Vesicle-like Virion of *Haloarcula hispanica* Pleomorphic Virus 3 Preserves High Infectivity in Saturated Salt. *Virology* **2016**, *499*, 40–51. [CrossRef]

41. Lee, S.T.M.; Ding, J.; Chiang, P.; Dyall-Smith, M.; Tang, S. Insights into Gene Regulation of the Halovirus His2 Infecting *Haloarcula hispanica*. *MicrobiologyOpen* **2020**, *9*, e1016. [CrossRef]

42. Bath, C.; Cukalac, T.; Porter, K.; Dyall-Smith, M.L. His1 and His2 Are Distantly Related, Spindle-Shaped Haloviruses Belonging to the Novel Virus Group, *Salterprovirus*. *Virology* **2006**, *350*, 228–239. [CrossRef] [PubMed]

43. Pietilä, M.K.; Laurinavičius, S.; Sund, J.; Roine, E.; Bamford, D.H. The Single-Stranded DNA Genome of Novel Archaeal Virus *Halorubrum* Pleomorphic Virus 1 Is Enclosed in the Envelope Decorated with Glycoprotein Spikes. *J. Virol.* **2010**, *84*, 788–798. [CrossRef] [PubMed]

44. Holmes, M.L.; Pfeifer, F.; Dyall-Smith, M.L. Analysis of the Halobacterial Plasmid pHK2 Minimal Replicon. *Gene* **1995**, *153*, 117–121. [CrossRef] [PubMed]

45. Thoden, J.B.; Holden, H.M. Active Site Geometry of Glucose-1-Phosphate Uridylyltransferase. *Protein Sci.* **2007**, *16*, 1379–1388. [CrossRef]

46. Alonso, M.D.; Lomako, J.; Lomako, W.M.; Whelan, W.J. A New Look at the Biogenesis of Glycogen. *FASEB J.* **1995**, *9*, 1126–1137. [CrossRef] [PubMed]

47. Holden, H.M.; Rayment, I.; Thoden, J.B. Structure and Function of Enzymes of the Leloir Pathway for Galactose Metabolism. *J. Biol. Chem.* **2003**, *278*, 43885–43888. [CrossRef] [PubMed]

48. Jaroentomeechai, T.; Stark, J.C.; Natarajan, A.; Glasscock, C.J.; Yates, L.E.; Hsu, K.J.; Mrksich, M.; Jewett, M.C.; DeLisa, M.P. Single-Pot Glycoprotein Biosynthesis Using a Cell-Free Transcription-Translation System Enriched with Glycosylation Machinery. *Nat. Commun.* **2018**, *9*, 2686. [CrossRef] [PubMed]

49. Egger, S.; Chaikuad, A.; Kavanagh, K.L.; Oppermann, U.; Nidetzky, B. UDP-Glucose Dehydrogenase: Structure and Function of a Potential Drug Target. *Biochem. Soc. Trans.* **2010**, *38*, 1378–1385. [CrossRef] [PubMed]

50. Dueber, E.L.C.; Corn, J.E.; Bell, S.D.; Berger, J.M. Replication Origin Recognition and Deformation by a Heterodimeric Archaeal Orc1 Complex. *Science* **2007**, *317*, 1210–1213. [CrossRef]

51. Duderstadt, K.E.; Chuang, K.; Berger, J.M. DNA Stretching by Bacterial Initiators Promotes Replication Origin Opening. *Nature* **2011**, *478*, 209–213. [CrossRef]

52. Takeda, S.N.; Nakagawa, R.; Okazaki, S.; Hirano, H.; Kobayashi, K.; Kusakizako, T.; Nishizawa, T.; Yamashita, K.; Nishimasu, H.; Nureki, O. Structure of the Miniature Type V-F CRISPR-Cas Effector Enzyme. *Mol. Cell* **2021**, *81*, 558–570.e3. [CrossRef]

53. Xiao, R.; Li, Z.; Wang, S.; Han, R.; Chang, L. Structural Basis for Substrate Recognition and Cleavage by the Dimerization-Dependent CRISPR–Cas12f Nuclease. *Nucleic Acids Res.* **2021**, *49*, 4120–4128. [CrossRef] [PubMed]

54. Harrington, L.B.; Burstein, D.; Chen, J.S.; Paez-Espino, D.; Ma, E.; Witte, I.P.; Cofsky, J.C.; Kyrpides, N.C.; Banfield, J.F.; Doudna, J.A. Programmed DNA Destruction by Miniature CRISPR-Cas14 Enzymes. *Science* **2018**, *362*, 839–842. [CrossRef]

55. Karvelis, T.; Bigelyte, G.; Young, J.K.; Hou, Z.; Zedaveinyte, R.; Budre, K.; Paulraj, S.; Djukanovic, V.; Gasior, S.; Silanskas, A.; et al. PAM Recognition by Miniature CRISPR–Cas12f Nucleases Triggers Programmable Double-Stranded DNA Target Cleavage. *Nucleic Acids Res.* **2020**, *48*, 5016–5023. [CrossRef] [PubMed]

56. Lee, H.H.; Yoon, J.Y.; Kim, H.S.; Kang, J.Y.; Kim, K.H.; Kim, D.J.; Ha, J.Y.; Mikami, B.; Yoon, H.J.; Suh, S.W. Crystal Structure of a Metal Ion-Bound IS200 Transposase. *J. Biol. Chem.* **2006**, *281*, 4261–4266. [CrossRef] [PubMed]

57. Muñoz-Espín, D.; Mateu, M.G.; Villar, L.; Marina, A.; Salas, M.; Meijer, W.J.J. Phage Φ29 DNA Replication Organizer Membrane Protein P16.7 Contains a Coiled Coil and a Dimeric, Homeodomain-Related, Functional Domain. *J. Biol. Chem.* **2004**, *279*, 50437–50445. [CrossRef] [PubMed]

58. Li, M.; Wang, R.; Zhao, D.; Xiang, H. Adaptation of the *Haloarcula hispanica* CRISPR-Cas System to a Purified Virus Strictly Requires a Priming Process. *Nucleic Acids Res.* **2014**, *42*, 2483–2492. [CrossRef] [PubMed]

59. Maier, L.-K.; Stachler, A.-E.; Brendel, J.; Stoll, B.; Fischer, S.; Haas, K.A.; Schwarz, T.S.; Alkhnbashi, O.S.; Sharma, K.; Urlaub, H.; et al. The Nuts and Bolts of the *Haloferax* CRISPR-Cas System I-B. *RNA Biol.* **2019**, *16*, 469–480. [CrossRef] [PubMed]

60. Koonin, E.V.; Makarova, K.S.; Wolf, Y.I. Evolutionary Genomics of Defense Systems in Archaea and Bacteria. *Annu. Rev. Microbiol.* **2017**, *71*, 233–261. [CrossRef]

61. Mercier, C.; Thies, D.; Zhong, L.; Raftery, M.J.; Erdmann, S. Characterization of an Archaeal Virus-Host System Reveals Massive Genomic Rearrangements in a Laboratory Strain. *Front. Microbiol.* **2023**, *14*, 1274068. [CrossRef]

62. Costa, A.R.; Van Den Berg, D.F.; Esser, J.Q.; Muralidharan, A.; Van Den Bossche, H.; Bonilla, B.E.; Van Der Steen, B.A.; Haagsma, A.C.; Fluit, A.C.; Nobrega, F.L.; et al. Accumulation of Defense Systems in Phage-Resistant Strains of *Pseudomonas aeruginosa*. *Sci. Adv.* **2024**, *10*, eadj0341. [CrossRef] [PubMed]

63. Kucukyildirim, S.; Ozdemirel, H.O.; Lynch, M. Similar Mutation Rates but Different Mutation Spectra in Moderate and Extremely Halophilic Archaea. *G3* **2023**, *13*, jkac303. [CrossRef] [PubMed]

64. Pfeiffer, F.; Losensky, G.; Marchfelder, A.; Habermann, B.; Dyall-Smith, M.L. Whole-Genome Comparison between the Type Strain of *Halobacterium salinarum* (DSM 3754 T) and the Laboratory Strains R1 and NRC-1. *MicrobiologyOpen* **2020**, *9*, e974. [CrossRef] [PubMed]

65. Pfeiffer, F.; Dyall-Smith, M. Genome Comparison Reveals That *Halobacterium salinarum* 63-R2 Is the Origin of the Twin Laboratory Strains NRC-1 and R1. *MicrobiologyOpen* **2023**, *12*, e1365. [CrossRef] [PubMed]