microLife, 2025, **6**, uqaf008
DOI: 10.1093/femsml/uqaf008
Advance access publication date: 19 May 2025
Research Article

Provirus deletion from *Haloferax volcanii* affects motility, stress resistance, and CRISPR RNA expression

Nadia Di Cianni[1], Simon Bolsinger[1], Jutta Brendel[1], Monika Raabe[2], Sabine König[2,3], Laura Mitchell[4], Thorsten Bischler[5], Tom Gräfenhan[5], Clarissa Read[6], Susanne Erdmann[7,8], Thorsten Allers[4], Paul Walther[6], Henning Urlaub[2,3], Mike Dyall-Smith[9,10], Friedhelm Pfeiffer[1,10], Anita Marchfelder[1,*]

[1]Molecular Biology and Biotechnology of Prokaryotes, Ulm University, Ulm 89069, Germany
[2]Max Planck Institute for Multidisciplinary Sciences, Bioanalytical Mass Spectrometry Group, Göttingen 37077, Germany
[3]Bioanalytics Groups, Department of Clinical Chemistry, University Medical Center Göttingen, Göttingen 37075, Germany
[4]School of Life Sciences, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, United Kingdom
[5]Core Unit Systems Medicine, University of Würzburg, Würzburg 97080, Germany
[6]Central Facility for Electron Microscopy, Ulm University, Ulm 89069, Germany
[7]Archaeal Virology, Max Planck Institute for Marine Microbiology, Bremen 28359, Germany
[8]Institute of Microbiology, University Innsbruck, Innsbruck 6020, Austria
[9]Veterinary Biosciences, Melbourne Veterinary School, Faculty of Science, University of Melbourne, Parkville 3010, Australia
[10]Computational Systems Biochemistry, Max Planck Institute of Biochemistry, Martinsried 82152, Germany
*Corresponding author. Molecular Biology and Biotechnology of Prokaryotes, Ulm University, 89069 Ulm, Germany. Email: anita.marchfelder@uni-ulm.de
Editor: [Axel Brakhage]

Abstract

Haloferax volcanii harbours four putative proviruses: Halfvol1, Halfvol2, Halfvol3, and Halfvol4. In this study, we successfully deleted all four provirus genomes, demonstrating, that they are not essential. Transcriptome comparison between this strain (ΔHalfvol1–4) and a wild-type strain reveals an increase in archaella and chemotaxis gene expression, resulting in higher swarming motility in ΔHalfvol1–4. Furthermore, ΔHalfvol1–4 cells show an elongated cell shape and a higher resistance to H_2O_2 stress compared to the wild type. RNA-seq also revealed downregulation of CRISPR arrays in the provirus-free strain. Circularised genomes of Halfvol1, Halfvol2, and Halfvol3 were found in the culture supernatant of the wild-type strain. This confirms excision of the proviruses from the chromosome, which seems to happen more efficiently at low temperature (30°C). Electron microscopy revealed potential viral particles in the supernatant, and mass spectrometry analysis confirmed the presence of structural viral proteins of Halfvol1 and Halfvol3 in the isolated virus sample. These observations suggest that these proviruses are active and cause a chronic infection in *H. volcanii*.

Keywords: *Haloferax volcanii*; archaea; provirus; virus; motility; CRISPR RNA

Introduction

Viruses are the most abundant entity on the planet and outnumber their microbial hosts by about a factor of 10 (Stern and Sorek 2011). Given their higher abundance, it is not surprising that cells are commonly found coinfected by multiple viruses, as indicated by many metagenomics studies (Roux et al. 2015). Viruses can exhibit different modalities of infection: virulent viruses follow a lytic cycle, resulting in cell lysis, while temperate viruses enter into a latent state, often by integrating their viral DNA into the host genome and being stably inherited. The integrated virus genome is referred to as a provirus and remains latent (or dormant) in a cell population unless induced to enter a lytic cycle. Spontaneous induction rates are usually very low, but sufficient to allow spread to new hosts. Additionally, some viruses exhibit a chronic cycle, where virus particles are continuously released without lysis of the host cell. Only a limited number of chronic viruses have been documented in prokaryotes, with the majority infecting archaea (Luke et al. 2014, Alarcon-Schumacher et al. 2022).

Some temperate viruses can integrate their genomes randomly into the host chromosome while others integrate in a site-specific manner, frequently into tRNAs (Campbell 1992). In the case of random integration, a gene may be targeted and disrupted. In the case of site-specific integration, the virus integrates into the host genome by recombining its viral attachment site (*attP*, P for phage) with the host's attachment site (*attB*, B for bacteria), catalysed by an integrase. As a result of this recombination, the provirus is inserted into the host DNA and flanked by two identical sequences, *attL* (Left) and *attR* (Right). Many archaeal and bacterial temperate viruses integrate into tRNA genes, using part of the tRNA as *attB* (Williams 2002, Badel et al. 2021). We here use the terms *attP* and *attB*, originally established for bacteriophages, also for the archaeal viruses.

Proviruses are surprisingly common in microorganisms (Casjens 2003, Roux et al. 2015, Touchon et al. 2017, Munson-McGee et al. 2018, McKerral et al. 2023), with ~60% of microorganism containing at least one functional or defective provirus in their genome (Casjens 2003). Viruses are generally perceived as enemies of their hosts, engaged in a perpetual arms race with microorganisms. However, proviruses may carry genes that are advantageous for the host itself. The merging of virus and host genomes, even if temporary, creates an ecological window for the evolution of mutually advantageous traits (Feiner et al. 2015). Lysogeny, thus may promote the development of symbiotic

Received 29 November 2024; **revised** 20 February 2025; **accepted** 15 May 2025

interactions: the long-term association of lysogenic viruses with their host may result in mutually beneficial interactions that enhance the reproductive success of both the virus and its host organism (Feiner et al. 2015). The provirus encoded superinfection exclusion mechanism is an example of such an advantageous trait, and is present in a wide variety of bacterial viruses, such as HK97, T4, D3, P1, and λ (Bondy-Denomy and Davidson 2014). Superinfection exclusion is a mechanism where a virus, which already has infected a host cell prevents subsequent infections by the same or other viruses (Folimonova 2012, Bondy-Denomy and Davidson 2014, Hunter and Fusco 2022). More specifically, superinfection exclusion occurs when a provirus blocks the adsorption (prophage P1), DNA injection (prophage Tuc2009), or DNA replication (prophage λ) of competing viruses, or when it harbours CRISPR-Cas loci targeting other viruses (Bondy-Denomy and Davidson 2014). Therefore, proviruses are no longer considered merely as parasites, but are also seen as mutualistic partners of the host (Bondy-Denomy and Davidson 2014).

Haloferax volcanii is a widely used model organism of archaea, with various aspects of its biology extensively studied, including replication, cell division, protein turnover, transcription, translation, and CRISPR-Cas system (Hartman et al. 2010, Leigh et al. 2011, Maier et al. 2019, Perez-Arnaiz et al. 2020, Schulze et al. 2020).

Notably, viruses are remarkably rare for the genus *Haloferax*, in contrast to other haloarchaeal genera. The first reported *Haloferax*-infecting virus was the HF1 virus infecting *Haloferax lucentense* and *H. volcanii* (Nuttall and Dyall-Smith 1993), however this virus is no longer available (M. Dyall-Smith, personal communication). Recently, a new siphovirus, Haloferax tailed virus 1 (HFTV1), was isolated, which infects *Haloferax gibbonsii* LR2-5 (Tittes et al. 2021a, Mizuno et al. 2019). Another recent addition is *H. volcanii* pleomorphic virus 1 (HFPV-1). This virus exhibits a persistent/chronic infection and is currently the sole available virus known to infect the model haloarchaeon *H. volcanii* (Alarcon-Schumacher et al. 2022).

Four proviruses have been identified in *H. volcanii* (Norais et al. 2007, Hartman et al. 2010, Dyall-Smith et al. 2021, Alarcon-Schumacher et al. 2022). All are encoded on the main chromosome. They are named Halfvol1, Halfvol2, Halfvol3, and Halfvol4. Proviruses Halfvol1 and Halfvol2 were shown to be active as they have been identified as circular genomes in a virus stock from *H. volcanii* (Dyall-Smith et al. 2021). Halfvol1 and Halfvol3 are distantly related and share characteristics which are typical for pleolipoviruses (Roine et al. 2010, Liu et al. 2015, Dyall-Smith et al. 2021). In general terms, pleolipoviruses are membrane vesicles with spike proteins, which encapsulate DNA genomes (Atanasova et al. 2015a, b, Demina and Oksanen 2020). Halfvol2 shows only limited similarity to known viruses and thus may represent a novel virus group (Dyall-Smith et al. 2021). The fourth provirus was described as a defective provirus (Norais et al. 2007, Hartman et al. 2010) and we denote it here as Halfvol4.

In order to further understand the importance of the proviruses for the host organism *H. volcanii*, we generated a strain that lacks all four proviruses from Halfvol1 to Halfvol4 (ΔHalfvol1–4) and characterised this provirus-free strain in detail.

Methods

Strains and growth conditions

Strains are listed in Table S3. *Haloferax volcanii* strain H100 was used as control strain for all experiments. In short, it is derived from wild-type strain DS2 by curing of plasmid pHV2 and deletion of three genes that can be used as selection markers (Δ*pyrE2*, Δ*leuB*, and Δ*hdrB*) (Allers et al. 2004). Strain H100 was grown aerobically with shaking (200 rpm) at 45°C in Hv-YPCab medium supplemented with thymidine. The parental strain of ΔHalfvol1–4 is H133 (Δ*pyrE2* Δ*leuB* Δ*hdrB* Δ*trpA*).

We use strain H100 (Δ*pyrE2* Δ*leuB* Δ*hdrB*) as a control which differs from strain H133 by carrying the *trpA* gene. This is consistent with the *trpA* gene being present in ΔHalfvol1–4 at the position of Halfvol3. However, while the *trpA* gene is under control of its native promoter in strain H100, it is under the control of p.*fdx* in ΔHalfvol1–4. Strain Δ*oapA* was previously described (Wolters et al. 2019, Mills et al. 2024), its parent strain is H26 (Δ*pyrE2*) (Allers et al. 2004). Strain ΔHalfvol1–4 was generated in the current study (see below).

Cloning of plasmids

Plasmids and primers are listed in Table S3. Plasmid pTA1102 [also termed pTA131-UP-DO(Halfvol4)] is a pTA131 plasmid with insertion of 1504 bp *H. volcanii* genomic DNA fragment containing downstream flanking region of Halfvol4 (bp 2 163 565–2 165 068), and 3100 bp *H. volcanii* genomic DNA fragment containing upstream flanking region of Halfvol4 (bp 2 108 260–2 111 359). The region deleted is bp 2 111 359–2 163 565, or HVO_2252–2293. To generate pTA131-UP-DO(Halfvol1), first the upstream region of Halfvol1 was amplified by polymerase chain reaction (PCR) using primers Provir6-up.fw and Provir6-up.rev. Plasmid pTA131 as well as the insert were digested with *Eco*RV and ligated together to generate pTA131-UP(Halfvol1). The downstream region of Halfvol1 was also amplified by PCR using primers ProVir6-do.fw and ProVir6-do.rev. The resulting fragment as well as plasmid pTA131-UP(Halfvol1) were digested with *Eco*RV and ligated to generate plasmid pTA131-UP-DO(Halfvol1). To generate pTA131-UP-DO(Halfvol2), the upstream region of Halfvol2 was amplified by PCR using primers Up KpnI Halvol2 F and Up HindIII Halvol2 R. Plasmid pTA131 and insert were digested with KpnI and HindIII and ligated together to generate pTA131-UP(Halfvol2). The downstream region of Halfvol2 was generated by PCR using primers Do HindIII Halvol2 F and Do XbaI Halvol2 R. The resulting fragment as well as plasmid pTA131-UP(Halfvol2) were digested with HindIII and XbaI and ligated together to generate plasmid pTA131-UP-DO(Halfvol2). pTA231-p.fdx-Nflag-bridgCDS Halfvol1 plasmid was obtained by amplifying the insert using primers CDS1-SnaBI F and CDS1-XbaI R. The resulting fragment and the pTA231-p.fdx-Nflag plasmid were subjected to digestion with SnaBI and XbaI, followed by ligation. pTA231-p.fdx-Nflag-bridgCDS Halfvol2 plasmid was obtained by amplifying the insert using primers CDS2-SnaBI F and CDS2-XbaI R. The resulting fragment and the pTA231-p.fdx-Nflag plasmid were subjected to digestion with SnaBI and XbaI, followed by ligation. pTA231-p.fdx-Nflag-bridgCDS Halfvol3 plasmid was obtained by amplifying the insert using primers CDS3-SnaBI F and CDS3-XbaI R. The resulting fragment and the pTA231-p.fdx-Nflag plasmid were subjected to digestion with SnaBI and XbaI, followed by ligation.

Deletion of proviruses

The provirus-free strain was generated using the pop-in pop-out method as previously described (Bitan-Banin et al. 2003). A Halfvol3 deletion strain was obtained from Uri Gophna (Tel Aviv University). In this strain, the Halfvol3 provirus is replaced by the *trpA* gene, which is controlled by the p.*fdx* promoter. The presence of the *trpA* marker at the position of Halfvol3 was confirmed by sequencing (Alarcon-Schumacher, personal

communication), but was inadvertently not reported in the original publication (Alarcon-Schumacher et al. 2022). The Halfvol4 locus was deleted in the ΔHalfvol3 strain using the deletion plasmid pTA1102 [pTA131-UP-DO(Halfvol4)]. Subsequently, Halfvol1 was deleted in the ΔHalfvol3ΔHalfvol4 strain using the plasmid pTA131-UP-DO(Halfvol1) and lastly, the provirus Halfvol2 was removed from the ΔHalfvol1ΔHalfvol3ΔHalfvol4 strain using the plasmid pTA131-UP-DO(Halfvol2). This resulted in strain ΔHalfvol1–4 which is devoid of all proviruses.

Determination of growth curves under normal and stress conditions

Growth curves were generated using 96-well microtiter plates and the Epoch 2 plate reader (BioTek Instruments). Cultures of strains H100 and ΔHalfvol1–4 were diluted to an OD_{650} of 0.05, and 200 μl of this diluted culture was added to the 96-well plates in triplicate or quintuplicate, along with a control containing only the Hv-YPCab supplemented thymidine medium. Media with 18% buffered salt water (BSW) (de Silva et al. 2021) were used for the standard salt condition, while 15% BSW and 23% BSW were used for low and high salt stress conditions, respectively. For each condition, at least three biological replicates and three technical replicates were performed. The standard temperature was 45°C, with 30°C and 50°C applied for temperature stress conditions. To minimize evaporation, 300 μl of medium was added to the outer wells of the plate. The optical density was recorded at 650 nm at 30-min or 1-h intervals (Dyall-Smith 2009). The growth data were analysed using Excel.

H_2O_2 survival assay

Haloferax volcanii was inoculated into 4 ml Hv-YPCab medium with thymidine and incubated at 45°C with shaking at 200 rpm. When the cultures reached an OD_{650} of 0.3–0.4, four 480 µl aliquots were taken and transferred into 2 ml reaction tubes. To these samples, H_2O_2 stock solution was added to achieve final concentrations of 2, 4, and 6 mM H_2O_2, resulting in 500 µl cell suspensions. For the control, 20 µl of 18% SW (Dyall-Smith 2009) was added instead of H_2O_2. Samples were incubated for 1 h at 45°C with shaking at 450 rpm. Serial dilutions were prepared using 18% SW, ranging from 10^{-3} to 10^{-6}, and 20 µl of each dilution was spotted in triplicate onto prewarmed Hv-YPC + thymidine agar plates and droplets were allowed to absorb. Plates were incubated for 3–4 days at 45°C. Colonies were counted, and survival fractions were calculated by averaging the triplicates for the 10^{-5} dilution level. Three biological replicates were performed.

UV survival assay

Cultures were grown to an OD_{650} of ∼0.4. Cells were then serially diluted (10^0–10^{-8}) in 18% BSW. Quadruplicate 20 μl samples were spotted onto Hv-YPC + thymidine plates, one plate for each exposure time (including a nonilluminated control). The plates were air dried for 30 min to allow the spotted culture to dry out. The plates were irradiated for 20, 40, 60, 80, and 100 s with a wave length of 254 nm. After irradiation, plates were incubated in a black bag at 45°C for 4 days. Following the incubation period, the colonies were counted. Three biological replicates were performed.

Light microscopy

Cultures with OD_{650} values of 0.3 (early exponential phase), 0.6 (late exponential phase), and 1.0 (stationary phase) were used. Cells were concentrated by centrifuging 1 ml of preculture at 4 000 × *g* for 6 min at room temperature. The supernatant was discarded, and the cell pellet was resuspended in 100–200 μl Hv-YPCab + thymidine medium. For microscopy, slides were prepared with a thin agar layer by boiling 0.1 g of agarose with 4 ml of ddH_2O and 6 ml of 30% saltwater, then mixing the solutions. A thin agarose bed was applied to a slide with a brush, briefly dried, and 3–5 μl of cell suspension was added. The sample was covered with a coverslip and sealed with clear nail polish. Microscopy was performed using a Leica DM5500 B light microscope in phase contrast at 100x magnification with oil immersion. At least three biological replicates were performed.

Motility analysis

To create soft agar plates (0.33%), 1 g of microagar was dissolved in 100 ml of demineralised water. After fully dissolving the agar in a microwave oven, 200 ml of prewarmed 30% BSW was added with gentle swirling. The mixture was autoclaved for sterilization and then placed on a heated magnetic stirrer plate, where 33 ml of 10x Hv-Ca and and other components ($CaCl_2$, uracil, thymidine, thiamine and biotin) were added as required. The liquid medium was poured into large plates, allowed to solidify for 1–2 h, and cultures were inoculated deeply into the semiliquid medium with a toothpick. Three replicates for each strain were spotted in the same plate. Plates were incubated upside down at 45°C for 4–5 days to observe swarming behaviour. Five biological replicates were performed.

Biofilm formation

Cultures were grown until late exponential phase, and 150 µl were added to each well of a microtiter plate, with 5–10 replicates and a negative control. Plates were incubated without shaking at 45°C and 30°C for 72 h. After incubation, liquid culture was removed, wells were washed twice with 18% BSW, fixed with 2% acetic acid, and stained with 0.1% crystal violet. Plates were washed three times with ddH_2O and air-dried. Finally, wells were destained with a solution of 10% acetic acid and 30% ethanol, and OD_{600} was measured in a microtiter plate photometer. Three biological replicates were performed.

RNA-seq analysis

Three biological replicates of the control and deletion strain cultures were incubated at 45°C and 200 rpm until 0.5 OD_{650} was reached. Then, 20 ml of culture was centrifuged at 4°C and 6000 × *g* for 10 min. The supernatant was discarded and the cell pellet was stored at −80°C until further use. For RNA preparation, the thawed pellet was resuspended in 1 ml NucleoZOL™ (Macherey-Nagel) and after the addition of 400 µl ddH_2O, the mixture was vortexed for 30 s and incubated at room temperature for 10 min to lyse the cells. After centrifugation at 13 000 × *g* for 15 min at room temperature, the upper phase was carefully transferred to a new tube and 5 µl of 4-bromanisol were added, followed by vortexing and incubation for 10 min. The tubes were centrifuged at 13 000 × *g* for 15 min at room temperature. The supernatant was transferred to a new tube and mixed with 1 ml of isopropanol. After incubation for 10 min at room temperature, the RNA was pelleted by centrifugation at 14 000 rpm for 10 min at 4°C. The pellet was washed twice with 70% ethanol and then resuspended in 240 µl of RNase-free ddH_2O.

To remove DNA, 30 µl of RQ1 DNase buffer and 30 µl of RNase-free RQ1 DNase were added and the mixture was incubated at 37°C for 1 h. RNA was then precipitated by ethanol and resuspended in 100 µl ddH_2O. Since *H. volcanii* is a polyploid organism with a large amount of DNA, the samples were subjected to

another step of DNase digestion using the TURBO™ DNase digestion according to the manufacturer's instructions. After 30 min incubation at 37°C, phenol–chloroform extraction was performed followed by ethanol precipitation. The RNA pellet was dissolved in 50 µl of RNase-free ddH_2O and the concentration determined using a NanoPhotometer® N60 (Implen). 10 µg of RNA samples were further processed at the Core Unit Systems Medicine facility at the University of Würzburg. Samples underwent depletion of ribosomal RNA molecules using a commercial Pan-Archaea riboPOOL rRNA depletion kit (siTOOLs Biotech, dp-K024-000027). The ribo-depleted RNA samples were first fragmented using ultrasound (4 pulses of 30 s at 4°C). Then, an oligonucleotide adapter was ligated to the 3' end of the RNA molecules. First-strand cDNA synthesis was performed using M-MLV reverse transcriptase with the 3' adapter as primer. After purification, the 5' Illumina TruSeq sequencing adapter was ligated to the 3' end of the antisense cDNA. The resulting cDNA was PCR-amplified using a high-fidelity DNA polymerase and the barcoded TruSeq-libraries were pooled in approximately equimolar amounts. Sequencing of pooled libraries, spiked with PhiX control library, was performed at 13–16 million reads per sample in single-end mode with 100 cycles on the NextSeq 2000 platform (Illumina). Demultiplexed FASTQ files were generated with bcl-convert v4.0.3 (Illumina). Raw sequencing reads were subjected to quality and adapter trimming via Cutadapt (Martin 2011) v2.5 using a cut-off Phred score of 20 and discarding reads without any remaining bases (parameters: –nextseq-trim=20 -m 1 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC). Afterwards, all reads longer than 11 nt were aligned to the *H. volcanii* DS2 reference genome [RefSeq assembly accession: GCF_000025685.1 excluding plasmid pHV2 (NC_013965.1)] using the pipeline READemption (Förstner et al. 2014) 2.0.3 with segemehl v0.3.4 (Hoffmann et al. 2009) and an accuracy cut-off of 95% (parameters: -l 12 -a 95). READemption gene_quanti was applied to quantify aligned reads overlapping genomic features by at least 10 nt (-o 10) on the sense strand based on annotations (CDS, mature_transcript, rRNA, and tRNA) originating from a detailed continuous manual curation effort (Laass et al. 2019). The annotation version from 28-MAR-2023 was used (annotated protein sequences in: https://doi.org/10.5281/zenodo.7794769), this version includes small ORFs reported in (Hadjeras et al. 2023). Based on these counts, differential expression analysis was conducted via DESeq2 (Love et al. 2014) v1.24.0. Read counts were normalised by DESeq2 and fold-change shrinkage was conducted by setting the parameter betaPrior to TRUE. In addition, outlier flagging by Cook's distance was disabled in the results function (cooksCutoff=FALSE).

Methods used for virus purification and concentration

Initial tests including high and low temperature (30°C and 50°C) and high and low salt concentrations (15% and 23%) revealed that provirus excision occurred at low temperature (30°C). 1 l cultures of strains Δ*oapA* and ΔHalfvol1–4 with an OD_{650} of 0.05 were incubated at 30°C with shaking at 180 rpm in Hv-YPCab and Hv-YPCab medium supplemented with thymidine until reaching the stationary phase (OD_{650} 0.86–1.0). The culture was centrifuged at 5524 × *g* at 4°C for 30 min. The supernatant, potentially containing viral particles, was subjected to another centrifugation step to eliminate any remaining cells. To precipitate the virus, 1:4 volume of cold 40% PEG6000 was added to the supernatant (final concentration 10%). After overnight incubation at 4°C, the viral suspension was centrifuged at 13 500 × *g* for 1 h at 4°C, and the pellet was resuspended in 10 ml of 18% BSW. Further purification involved three filtration steps: the first with a 0.45 µm pore-size filter, followed by two with 0.22 µm filters. Another precipitation step with cold 40% PEG6000 was performed, followed by overnight incubation at 4°C. After centrifugation at 15 000 × *g* for 1 h at 4°C, the supernatant was removed, and the pellet was resuspended in 1 ml of 18% BSW. To eliminate free nucleic acids, 10 µl of RNase A and 20 µl of RQ1 RNase-free DNase (1 U/µl, Promega) were added to the sample, which was then incubated for 2 h at 37°C. The digested samples were either stored at 4°C or immediately processed for CsCl density-gradient centrifugation.

A CsCl solution, having a density of 1.2 g/ml and a concentration of 0.342 g/ml in 18% BSW, was filtered through a 0.22 µm filter. Subsequently, 11 ml of this solution was dispensed into Ultra-Clear™ centrifuge tubes (14 mm × 95 mm), followed by gently layering 1 ml of the viral sample over the CsCl solution. The tube was centrifuged at 38 000 × *g* for 21-24 h at 4°C using the SW40T1 rotor in the Optima L-60+ Ultracentrifuge (Beckman Coulter). After ultracentrifugation, the bands were collected by carefully aspiring 1 ml fractions of the gradient from the top. Each fraction was supplemented with 10% PEG6000 (final concentration) and incubated overnight at 4°C. Following centrifugation at 15 000 × *g* for 1 h at 4°C, the pellet was resuspended in 50–100 µl of 18% BSW. To analyse for the presence of viral genomes, PCR analysis was conducted using 1 µl of a 1:10 diluted sample as the template. A total of 30 or 31 cycles was applied.

Detection of provirus excision by PCR analysis

To detect virus DNA circularization, a PCR was performed using divergent primers located at the ends of the provirus region (illustrated in Fig. 5). The primers employed for the detection of circular Halfvol1 are prov6 circ R and prov6 circ F; for Halfvol2, Prov2 circ F and Prov2 circ R were used; and for Halfvol3, prov5 circ F and prov5 circ R were employed (Table S3C.). The PCR mixture included 5 µl of 5x Green GoTaq Reaction Buffer, 1.25 µl of DMSO, 0.5 µl of dNTPs (5 mM), 0.25 µl each primer (500 ng/µl), 1 µl of template (a 1:10 dilution of isolated virus particles from the culture supernatant to avoid salt interference), 17.5 µl of ddH_2O and 0.25 µl of Go-Taq polymerase, for a total reaction volume of 25 µl. The initial denaturation was carried out at 95°C for 5 min, followed by 31 cycles of denaturation at 95°C for 30 s, annealing temperature at 62°C for 30 s, and extension at 72°C for 30 s. After cycling, a final extension was performed at 72°C for 3 min. PCR products were analysed on a 1% agarose gel. The resulting bands were excised and sent for Sanger sequencing.

Transmission electron microscopy

Samples from culture supernatants were fixed with 2.5% glutaraldehyde in 18% BSW for 25 min. Approximately 7 µl of the fixed samples were placed onto glow-discharged 300 mesh copper grids covered with a carbon-enhanced formvar film. After 10 min, excess liquid was removed by tapping the grid on filter paper soaked with ddH_2O. The grids were then washed three times in droplets of ddH_2O, with excess water removed after each wash. The sample was stained with a 2% aqueous uranyl acetate solution for 4 s. After drying, the samples were imaged using transmission electron microscopy (TEM; JEM1400, Jeol, Japan).

Mass spectrometry

The second band derived from strain Δ*oapA* and the corresponding pink band from ΔHalfvol1–4 after CsCl density-gradient centrifugation were concentrated by polyethylenglykol (PEG)

Table 1. Proviruses of *H. volcanii*. This table lists the key characteristics of the four proviruses, which have been reported for *H. volcanii*. A complete list of their characteristics is shown in Table S1. All four proviruses contain direct repeats of 14 nt. For Halfvol1-3, only the *attP* is contained, thus excluding the repeat copy from *attB*. Thus, the provided length represents the circularised (activated) form of the provirus. The morphotype 'novel group' is from Dyall-Smith et al. (Liu et al. 2015, Dyall-Smith et al. 2021).

Name	Size (bp)	Morphotype	Flanks (*attP*)	Number of integrases
Halfvol1	20 573	Beta-pleolipovirus	$tRNA^{Pro}$-TGG	1
Halfvol2	12 275	Novel group	$tRNA^{Ala}$-CGC	2
Halfvol3	12 527	Alpha-pleolipovirus	$tRNA^{Arg}$-TCG	1
Halfvol4	53 702	Not known	–	3

precipitation (see above). Three biological replicates were prepared. The pellet was resuspended in ∼100 µl of 50 mM Tris–HCl pH 7.2. The samples were centrifuged at 10 000 × *g* for 10 min to remove insoluble particles, and protein concentration was measured using a Bradford assay (Roti-Nanoquant, Carl Roth GmbH). 10 µg of protein from each sample was loaded onto a 12% sodium dodecyl sulfate (SDS) gel, which was stained with Coomassie Brilliant Blue. sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE)-separated protein samples were processed as described by Shevchenko et al. (1996a, b). The resulting peptides were loaded to nano High-performance liquid chromatography (HPLC) coupled to an Exploris Orbitrap Mass spectrometer (Thermo Fisher). The peptides were separated with a linear gradient of 5%–40% buffer B (80% acetonitrile and 0.1% formic acid) at flow rate of 300 nl/min over 52 min total gradient time. The MS data was acquired by scanning the precursors in mass range from 350 to 1400 m/z at a resolution of 60 000 at m/z 200. Top30 precursor ions were chosen for MS2 by using data-dependent acquisition mode at a resolution of 15 000 at m/z 200. Data analysis and search was performed using Maxquant Software (1.6.17.0) and Scaffold 5.2.1 with a manually curated Haloferax_volcanii database (release March 2023; 4222 entries; available under DOI:10.5281/zenodo.7794769) and a database of common contaminants (supplied with the Maxquant Software; version 1.6.17.0).

Pulsed-field gel electrophoresis

For pulsed-field gel electrophoresis (PFGE), genomic DNA was prepared in agarose plugs as described previously (Delmas et al. 2009). For analysis of intact genomic DNA, agarose plugs were subjected to 100 Gy of γ radiation using a ^{137}Cs source (Gammacell 1000), to linearize circular chromosomes (Beverley 1989). PFGE was performed using a CHEF Mapper apparatus (Bio-Rad). DNA fragments were separated on a 1.2% agarose gel in 0.5 × TBE (Tris, borate, and EDTA) at 14°C, with a gradient voltage of 6 V/cm, linear ramping, an included angle of 120°, initial and final switch times of 0.64 s and 1 min 13.22 s, respectively, and a run time of 20 h 46 min, as described previously (Ausiannikava et al. 2018). The gel was stained with ethidium bromide.

Quantitative PCR

For sample preparation, 1 ml of culture was grown to OD_{650} 0.8 and centrifuged at 1503 × *g* for 8 min at room temperature. The supernatant was filtered through a 0.22-µm filter to remove cells. Subsequently, 1 ml lysis buffer was added and ethanol precipitation was performed. After centrifugation, the pellet was resuspended in 300 µl of ddH_2O. Real-time quantitative PCR (RT-qPCR) was performed using a qtower3 touch qPCR machine (Analytik Jena) in 96-well plate. The primers used for each target were qPCR Halfvol1 F2 and qPCR Halfvol1 R2 for Halfvol1; qPCR Halfvol2 F1 and qPCR Halfvol2 R1 for Halfvol2; and qPCR Halfvol3 F2 and qPCR Halfvol3 R2 for Halfvol3 (Table S3C). The qPCR mixture included 10 µl of qPCRBIO SyGreen Blue Mix Separate-ROX (PCR Biosystems), 0.8 µl each primer (10 µM), 6.4 µl of ddH_2O, and 2 µl of template (1:2 or 1:10 diluted sample), for a total reaction volume of 20 µl. The initial denaturation was carried out at 95°C for 2 min, followed by 40 cycles of denaturation at 95°C for 5 s, annealing/extension temperature at 60°C for 30 s. First primer efficiencies were assessed (Halfvol1: 0.88 Halfvol2: 0.98 Halfvol3: 0.87). The level of circularised viral genome was quantified using the absolute quantitative method (Harshitha and Arunraj 2021). Standard curves were generated with plasmid pTA231-p.fdx-Nflag-bridgCDS Halfvol1, pTA231-p.fdx-Nflag-bridgCDS Halfvol2, and pTA231-p.fdx-Nflag-bridgCDS Halfvol3 to calculate the absolute concentration of circularised Halfvol1, Halfvol2, and Halfvol3, respectively (Table S3B.). Each sample was analysed in triplicate for both technical and biological replicates.

Results

Proviruses of *H. volcanii*

The genome of the *H. volcanii* type strain DS2 contains four predicted proviruses, and their key characteristics are given in Table 1 and Table S1. The genomic organization of Halfvol1, Halfvol2, and Halfvol3 is illustrated in Fig. S1. Halfvol1, Halfvol2, and Halfvol3 are predicted to be double-stranded DNA viruses, each of them flanked by a perfect direct repeat (14 nt long), or *att* site. This repeat is tRNA-derived (all 14 nt for Halfvol2 and 13 of the 14 nt for Halfvol1 and Halfvol3). At one end is a complete tRNA gene and at the opposite end is a partial copy of that tRNA.

As was reported previously (Roine et al. 2010, Liu et al. 2015, Dyall-Smith et al. 2021), Halfvol1 and Halfvol3 encode proteins, which are homologous to those from pleolipoviruses (Table 2). In particular, the structural proteins VP4 (spike protein) and VP3 are present in Halfvol1 (HVO_0271; HVO_0269) and in Halfvol3 (HVO_1431; HVO_1432). Moreover, genes homologous to VP8-like NTPases (HVO_0274; HVO_1426; and HVO_1428) are present in both proviruses. In addition, both proviruses encode winged-HTH domain proteins (HVO_0260 and HVO_0261; HVO_1423) with distant similarity to the halovirus phiH repressor protein which is important for lysogeny (Ken and Hackett 1991).

Attempts to identify a Halfvol2 protein related to a virus capsid protein were not successful, as sequence similarity searches

Table 2. Proteins encoded on proviruses Halfvol1 and Halfvol3, which are related to key virus proteins. Proteins which are encoded on proviruses Halfvol1 and Halfvol3 and which are homologous to viral proteins are shown. Excluded are (A) integrases and (B) HTH domain proteins distantly related to the halovirus phiH repressor. No proteins homologous to virus proteins are encoded on Halfvol2 and Halfvol4. Virion protein indicates proteins present in purified virus particles. Spike protein is the viral attachment protein. The term VP8-like ATPase is used for homologues of HRPV1-VP8 (Pietilä et al. 2016, Wang et al. 2018). VP serial numbers are not stable between different viruses.

Gene	Pleolipovirus homologues	Identity (%)	Function	References
	Halfvol1. All proteins are specified by a conserved gene cluster in beta-pleolipoviruses.			
HVO_0268	HGPV1-VP2	71	Virion internal membrane protein	Atanasova et al. (2018)
HVO_0269	HFPV1-VP3	80	Virion internal membrane protein	Alarcon-Schumacher et al. (2022)
HVO_0271	HGPV1-VP4	50	Virion spike protein	Atanasova et al. (2018)
HVO_0272	HFPV1-VP5	77	Virion protein	Alarcon-Schumacher et al. (2022)
	HGPV1-VP5	33		
HVO_0273	HFPV1-VP6	77	Virion protein	Alarcon-Schumacher et al. (2022)
	HGPV1-VP6	52		
HVO_0274	HGPV1-VP7	54	VP8-like ATPase	Atanasova et al. (2018)
	HFPV1-VP7	52		
	Halfvol3. All but HVO_1434 are specified by a conserved gene cluster in alpha-pleolipoviruses.			
HVO_1426	HRPV1-VP8	53	VP8-like ATPase	Demina and Oksanen (2020)
HVO_1428	HHPV1-VP6	40	VP8-like ATPase	Demina and Oksanen (2020)
	HRPV1-VP7	35		
HVO_1429	HRPV2-VP6	32	Conserved protein	Demina and Oksanen (2020)
HVO_1431	HRPV1-VP4	42	Virion spike protein	Demina and Oksanen (2020)
	Hardyhisp2-VP1	32		
HVO_1432	HHPV2-VP3	63	Virion internal membrane protein	Demina and Oksanen (2020)
	HRPV1-VP3	53		
HVO_1434	HRPV2-VP1	40	Putative replicase	Demina and Oksanen (2020)
	HRPV6-VP1	40		

(BLASTp) did not find matches to known capsid proteins. Also, a foldseek analysis did not reveal any capsid protein showing significant (better than E-value < 10^{-3}) structural similarity. However, a packaging ATPase candidate could be retrieved. HVO_0377 shows structural similarity to the proposed packaging ATPases from *Haloarcula hispanica* virus PH1 (E-value < 10^{-10}; HhPH1_gp13; UniProt: M4JFA3) and from Sulfolobus turreted icosahedral virus 2 (E-value < 10^{-11}; STIV2_B204; UniProt: D5IEZ9; PDB:4KFU).

Halfvol4 is characterised by high A/T content when compared to the main chromosome (Halfvol4 has a 50.1% GC content while the rest of the genome has a 66.7% GC content) and it displays an unusual codon usage variation, when compared to the codon usage of the whole chromosome (Hartman et al. 2010). Halfvol4 contains three XerC/D-like integrases, a type IV secretion system (T4SS), two divergent *orc* genes, a restriction modification system (Hartman et al. 2010), and the recently identified new defence system Avast (HVO_2276, Avast type 2) (Gao et al. 2020, Payne et al. 2022). The Halfvol4 region was previously reported as a defective provirus (Norais et al. 2007, Hartman et al. 2010). Initially, this provirus was attributed to genome positions 2110297–2163576 (covering HVO_2252 to HVO_2293A) (Hartman et al. 2010) or covering HVO_2252 to HVO_2293 (Norais et al. 2007). We shifted the boundaries of Halfvol4 to genome positions 2109865–2163566 (covering HVO_2251A to HVO_2293A), after finding that this element is enclosed by a 14-nt perfect direct repeat. A third perfect copy and a near-complete copy (13 nt) of this repeat is located internally within Halfvol4.

In some closely related haloarchaeal species, this region shows heterogeneity (Table S2). While in *H. volcanii* Halfvol4 is inserted at this site, in other haloarchaeal species there is no insertion present, while still others carry an insert, but of different size (Table S2). The genomic organization of Halfvol4 is illustrated in Fig. S1.

Generation of the provirus-free *H. volcanii* strain ΔHalfvol1–4

To investigate the roles and impacts of proviruses in *H. volcanii*, we generated a deletion strain in which all four predicted proviruses were removed (ΔHalfvol1–4). In a previous study, the provirus Halfvol3 was deleted from strain H133 (Δ*pyrE2* Δ*leuB* Δ*hdrB* Δ*trpA*) and substituted with the *trpA* marker (Alarcon-Schumacher et al. 2022). From this ΔHalfvol3 strain, the Halfvol4 locus was removed, followed by the deletion of Halfvol1 from the ΔHalfvol3ΔHalfvol4 strain. Lastly, Halfvol2 was deleted from the ΔHalfvol1ΔHalfvol3ΔHalfvol4 strain resulting in the provirus-free strain (ΔHalfvol1–4), which was confirmed by Southern analysis (Fig. S2).

For the characterisation of the provirus-free strain (ΔHalfvol1–4), strain H100 (Δ*pyrE2* Δ*leuB* Δ*hdrB*) instead of the parent strain H133 was used as a control since H100 contains the same genetic markers as ΔHalfvol1–4 (Table S3). This is due to the fact that Halfvol3 was replaced by the *trpA* gene (under the p.fdx promoter). Previous findings indicate that, in the *H. volcanii* DS70 strain, the pHV4 mini-chromosome was integrated into the main chromosome through recombination between the ISHvo9 transposons coding for transposase HVO_0278 (chromosome) and HVO_A0279 (pHV4) (Hawkins et al. 2013). Given that HVO_0278 is encoded within the Halfvol1 locus, this suggests that pHV4 is integrated into Halfvol1. Since pHV4 carries multiple essential

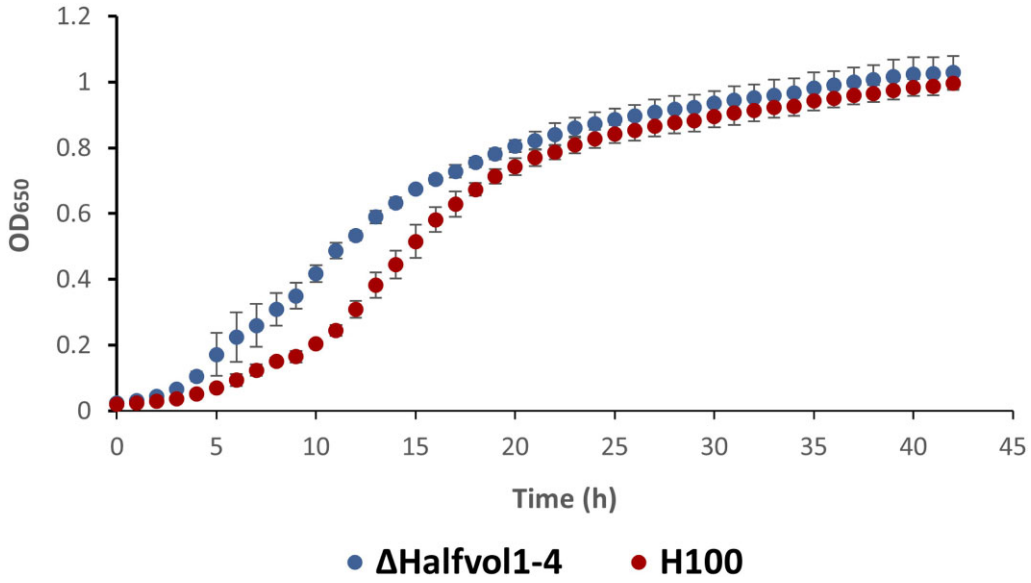

Figure 1. Growth curves of provirus-free strain and control strain. Both strains were grown under standard conditions (45°C and 18% BSW). The provirus containing control is *H. volcanii* strain H100 (red) and its growth has been compared with that of the provirus-free strain ΔHalfvol1–4 (blue). Data shown are based on three biological replicates each, see the section 'Methods' for details.

genes (Norais et al. 2007), Halfvol1 can only be deleted if pHV4 becomes episomal again. PFGE (Fig. S3) revealed the excision of mini-chromosome pHV4 from the chromosome in ΔHalfvol1–4. Because the Halfvol1 locus, and therefore the site of the integration of pHV4, has been deleted in this strain, pHV4 exists exclusively in its episomal form in the provirus-free strain.

Phenotypic characterisation of the provirus-free strain ΔHalfvol1–4

A number of phenotypic tests were used to compare the ΔHalfvol1–4 strain with the isogenic H100 strain. Growth analysis at optimal temperature and salt concentration (45°C and 18% BSW) indicates that the ΔHalfvol1–4 strain has a slightly shorter lag phase than the control strain (Fig. 1). Growth curves were also tested under salt stress (15% and 23% BSW) and temperature stress (30°C and 50°C). Under low salt conditions and both low and high temperatures, the provirus-free strain is growing slightly better than the control (Fig. S4).

Phase-contrast microscopy analysis revealed that the ΔHalfvol1–4 strain predominantly exhibited an elongated cell shape, in contrast to the control H100 strain, which displayed a round cell shape. However, round-shaped cells were also observed in ΔHalfvol1–4, primarily during the late exponential phase (OD_{650}: 0.6) (Fig. 2). Viruses and proviruses can significantly influence the motility of their host organisms (Wang et al. 2009, Brathwaite et al. 2015). To explore this, a swarming assay was conducted with ΔHalfvol1–4 and a control strain (H100). The ΔHalfvol1–4 strain exhibited hypermotility compared to the control, suggesting that the presence of proviruses reduces motility (Fig. 3). In addition, biofilm formation was investigated but no significant difference was observed between the provirus-free strain and the control. (Fig. S5). The effects of provirus deletion on cell responses to stressors were investigated in two experiments: oxidative stress and UV survival assays. The ΔHalfvol1–4 strain demonstrated high resilience to oxidative stress at a high H_2O_2 concentration (6 mM), compared to the control (Fig. 4). However, no significant differences were observed in response to UV stress (Fig. S6).

Transcriptome analysis

To assess whether provirus removal affects gene expression, the transcriptome of the provirus deletion strain ΔHalfvol1–4 was compared to that of the control strain H100 by RNA-seq (Tables 3 and 4, Table S4). In total, 79 genes were upregulated and 233 genes were downregulated in the deletion strain, based on the criteria of log2 fold change >1 or < −1 and a *P*-adjusted value < .01. The RNA-seq data indicates an increased expression of archaella, chemotaxis and pili genes in the ΔHalfvol1–4 strain, which are involved in motility. Specifically, the locus from HVO_1203 to HVO_1225, which contains genes associated with motility, shows upregulation of ∼2- to 3-fold in ΔHalfvol1–4 (Table 4). This genetic expression pattern is consistent with the hypermotility observed in the swarming assay (Fig. 3), thereby providing a molecular explanation for the observed phenotypic behaviour. Moreover, in ΔHalfvol1–4, a pronounced downregulation of CRISPR arrays (P1, P2, and locus C) was observed, reaching up to a 9-fold reduction compared to the control strain (Table 3, Table S4). The reduction in crRNA expression in the provirus-free strain suggests an influence of proviruses on the adaptive immune system of *H. volcanii*.

Additionally, transcription regulators (HVO_A0394, HVO_A0135, HVO_C0076, HVO_A0527, and HVO_1052) and genes involved in the amino acid metabolism (HVO_A0295 and HVO_0285) also exhibited decreased expression levels (Table S4). A more comprehensive functional annotation analysis of differentially expressed genes in ΔHalfvol1–4 compared to H100 is provided in Fig. S7, utilizing the Database for Annotation, Visualization, and Integrated Discovery (DAVID) bioinformatics tool (Huang et al. 2009, Sherman et al. 2022). Interestingly, the DAVID analysis revealed that a notably high number of membrane-associated proteins are differentially expressed in the ΔHalfvol1–4 strain. However, motility-related genes appear to be underestimated and non-COG genes (such as CRISPR arrays, tRNAs, and ncRNAs) are excluded from the DAVID analysis as they lack COG classification. It should be noted that proteins with the code prefix HVO_A originate from mini-chromosome pHV4, which is integrated into the major chromosome in the

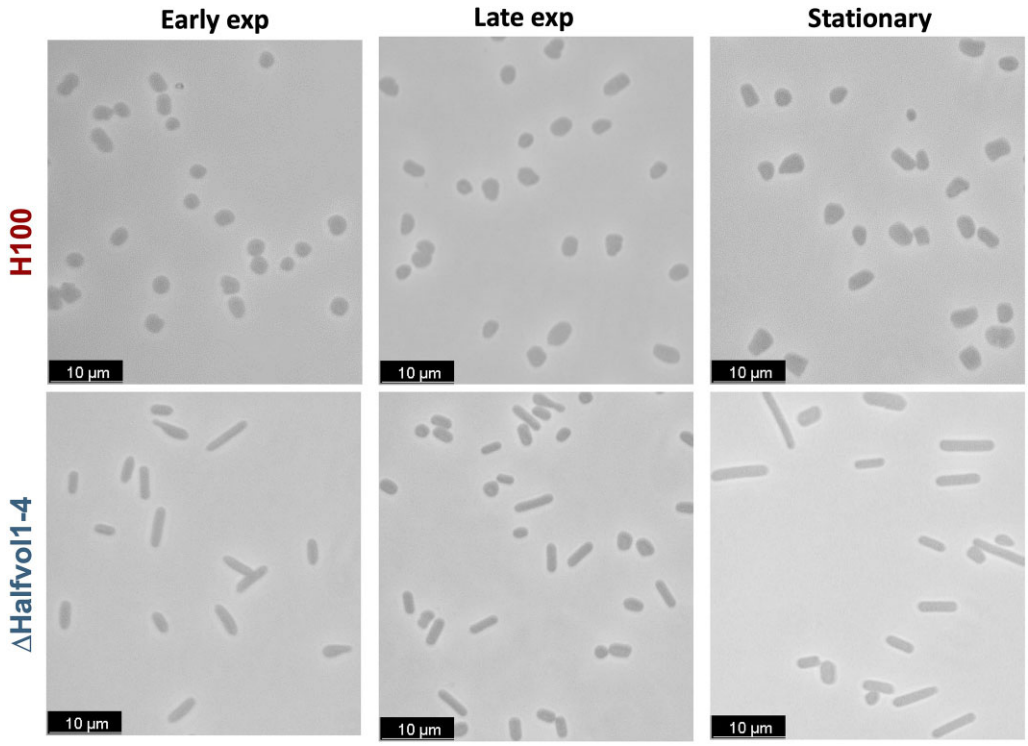

Figure 2. Microscopy analysis of provirus-free strain and control strain. Phase contrast microscopy images of the control strain H100 and the provirus-free strain ΔHalfvol1–4. The first column shows images of H100 and ΔHalfvol1–4 in early exponential phase (OD_{650} 0.3), the second column in late exponential phase (OD_{650} 0.6), and the third column in stationary phase (OD_{650} 1.0). The images are representative for three biological replicates, which were analysed for each strain.

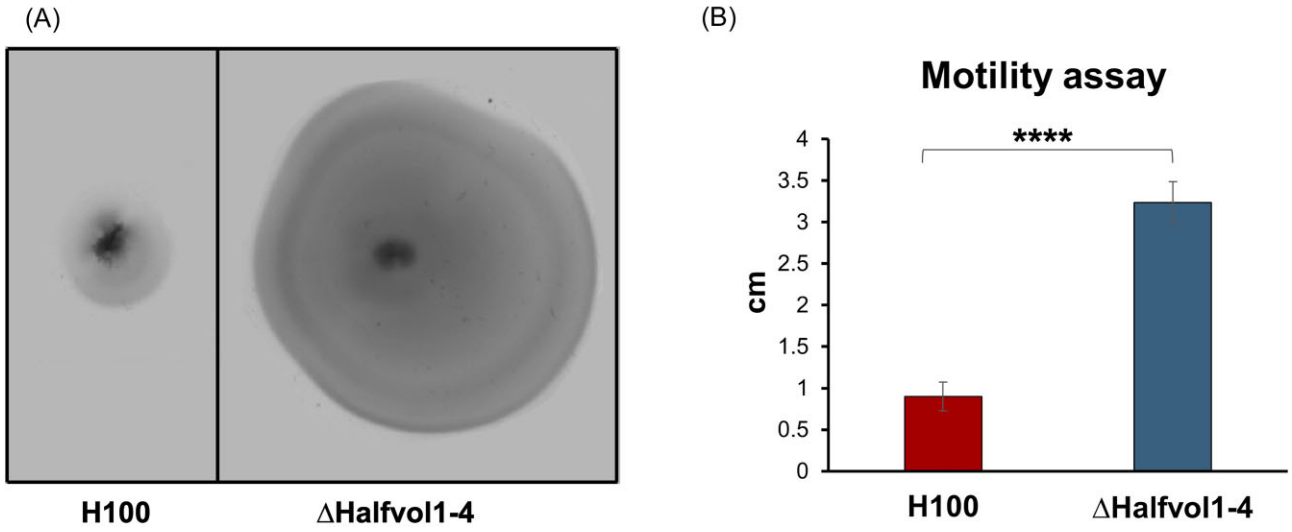

Figure 3. Swarming assay of provirus-free strain and control strain. Swarming motility of the control strain H100 and the provirus-free strain ΔHalfvol1–4 was tested at 45°C in Hv-Ca plates (0.3% agar). (A) Phenotypical observation of the halo disk due to the swarming of strains H100 and ΔHalfvol1–4. One of five biological replicas is shown. (B) The measured diameter of the halo disks. Values shown are based on five biological replicates. Asterisks indicate significant differences (*t*-test); ****: highly significant (*P*-value ≤ .0001).

parent strain but occurs in episomal form in strain ΔHalfvol1–4. Changes in expression of these genes might also be related to this. Curiously, several genes immediately downstream of Halfvol1 (HVO_0281–HVO_0289) were upregulated up to 8-fold (log2FC 3.1) in the ΔHalfvol1–4 deletion strain. This suggests that the deletion of Halfvol1 may lead to the upregulation of these genes, indicating a possible regulatory effect of Halfvol1 on its downstream region.

Proviruses excise from the chromosome and are released from cells

Two different approaches were used to detect provirus induction: (1) PCR and Sanger sequencing were used to analyse the *attP* and *attB* sites, to identify configurations indicative of provirus excision from the chromosome (in strains H100, H133, and Δ*oapA*, see Figs 5, 7, and 8B.). (2) Examination of culture supernatants by (a) electron microscopy to detect virus-like particles (Fig. 9 and

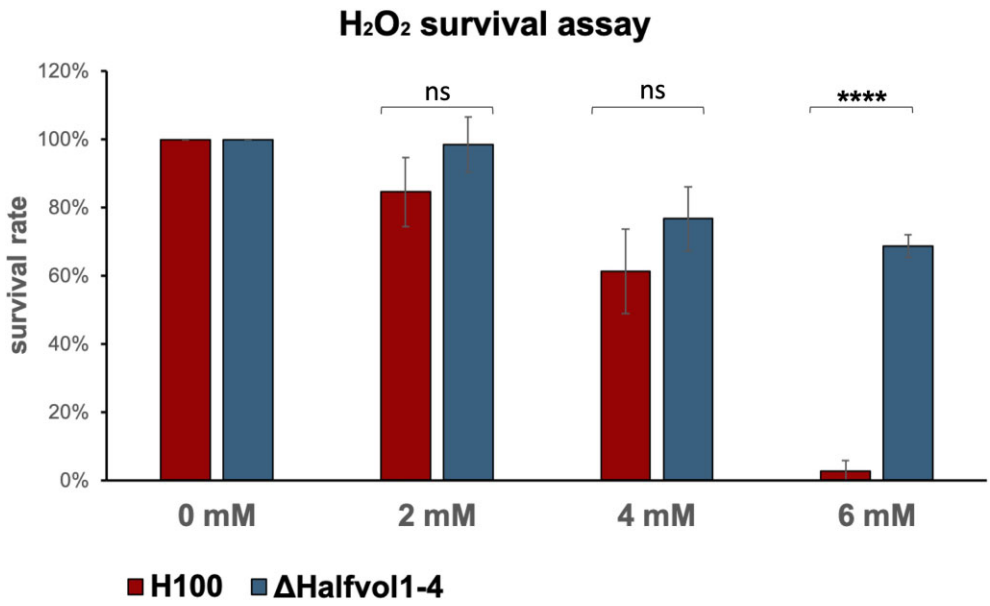

Figure 4. Oxidative stress resilience between provirus-free strain and control strain. The cultures were incubated for 1 h in 18% salt water with different concentrations of H_2O_2 (0, 2, 4, and 6 mM). Cells were then plated on Hv-YPC + thymidine agar plates and the survival rate of the strains was determined. The values shown are based on three biological replicates. Symbols shown on top of the bar: asterisks indicate significant differences (*t*-test); ****: highly significant (*P*-value ≤ .0001), ns = not significant.

Table 3. Differential gene expression between ΔHalfvol1–4 and strain H100 (control). Many genes are differentially regulated in the provirus deletion strain. In this table, only the 10 most strongly up- and downregulated genes are listed. Provirus genes that were deleted in ΔHalfvol1–4 and are therefore highly downregulated are not shown. The log2 fold change (column log2FC) deletion strain versus wild-type strain is given. Upregulation is indicated by positive log2FC values, downregulation by negative log2FC values. The HVO gene number (column Gene), gene annotation (column Annotation), and adjusted *P*-value (column padj) is also provided. Only genes with an adjusted *P*-value < 10^{-06} are included).

Gene	Annotation	log2FC	padj
	Top 10 upregulated genes		
HVO_0281	tRNA-Pro	3.08	2.5×10^{-109}
HVO_0288	tRNA-His	2.10	1.84×10^{-50}
HVO_0287	Conserved hypothetical protein	2.08	2.81×10^{-32}
HVO_0283	Archaea-specific helicase AshA	2.06	9.14×10^{-53}
HVO_0282	DASS family transport protein	2.05	6.19×10^{-25}
HVO_B0065	Thioredoxin domain protein	1.75	1.02×10^{-08}
HVO_B0053	DUF3209 family protein	1.63	1.24×10^{-07}
HVO_0284	Cupin 2 barrel domain protein	1.60	4.14×10^{-12}
HVO_0285	M50 family metalloprotease	1.60	7.41×10^{-13}
HVO_1228	DUF5059 domain/halocyanin domain protein, hcpE	1.55	1.66×10^{-07}
	Top 10 downregulated genes		
HVO_A0295	Amidase (hydantoinase/carbamoylase family, amaB2)	−3.53	6.35×10^{-21}
HVO_A0129s	Small RNA	−3.34	1.28×10^{-33}
CRISPR2	CRISPR Locus P1	−3.18	3.33×10^{-43}
HVO_A0394	HTH domain protein	−3.13	4.32×10^{-77}
HVO_1252	DUF7501 domain protein	−3.11	3.43×10^{-96}
HVO_A0212s	Small RNA	−3.07	1.9×10^{-31}
CRISPR3	CRISPR Locus P2	−2.93	7.79×10^{-31}
HVO_A0129	Conserved hypothetical protein	−2.69	2.3×10^{-18}
CRISPR1	CRISPR Locus C1	−2.61	1.05×10^{-31}
HVO_2307	HAD superfamily hydrolase	−2.59	1.75×10^{-77}

Table 4. Differential gene expression of motility genes. Motility related genes showing upregulation are listed. Genes are listed when at least 2-fold regulated (log2FC) > 1 and when the value is significant (adjusted *P*-value < .075).

Gene	Annotation	log2FC	padj
HVO_1211	Archaellin A2	1.69	0.0058
HVO_2220	Transducer protein Htr38	1.54	0.0011
HVO_2450	Pilin PilA3	1.48	0.0124
HVO_1205	Taxis cluster protein CheD	1.47	0.0182
HVO_1214	Arl cluster protein ArlF	1.32	0.0077
HVO_1225	Purine-binding taxis protein CheW1	1.31	0.0058
HVO_1203	Arl cluster protein ArlD	1.24	0.0521
HVO_1221	Taxis protein CheF1	1.18	0.0162
HVO_1217	Archaellar motor/biogenesis protein ArlI	1.17	0.0123
HVO_1222	Protein-glutamate O-methyltransferase CheR	1.14	0.0005
HVO_1215	Arl cluster protein ArlG	1.15	0.0740
HVO_1216	Arl cluster protein ArlH	1.13	0.0736
HVO_1206	Taxis cluster protein CheC	1.11	0.0187
HVO_1218	Archaellar motor/biogenesis protein ArlJ	1.08	0.0274
HVO_2451	Pilin PilA4	1.08	0.0002
HVO_0555	Transducer protein Htr15a	1.03	0.0016
HVO_1219	Taxis protein CheF2	1.02	0.0186

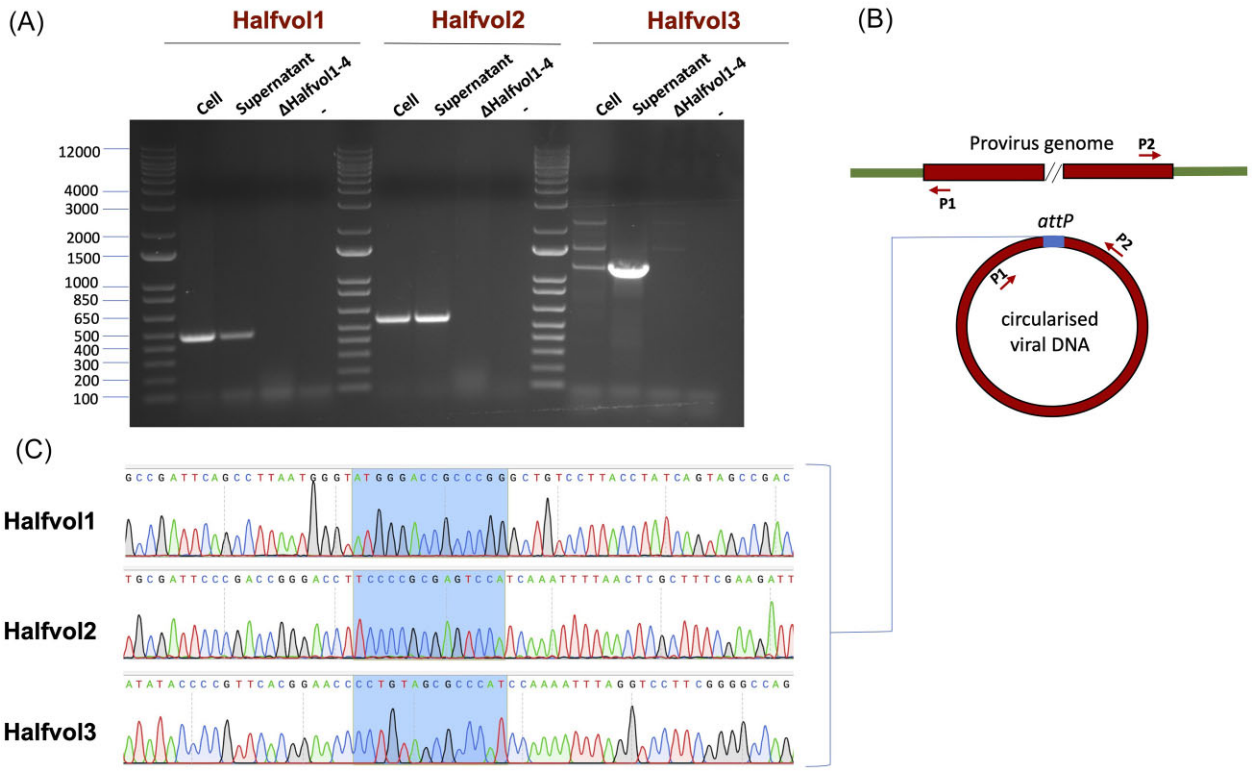

Figure 5. Detection of *attP* site of Halfvol1, Halfvol2, and Halfvol3, indicating provirus excision. (A) PCR products obtained with outward facing primers, indicating the circularization of Halfvol1, Halfvol2, and Halfvol3. Lane 'Cell': wild-type cells (H119) were taken from a plate and used as template. In 'Supernatant', 1 µl of 1:10 dilution of isolated excised proviruses from the culture supernatant (H133) from liquid cultures was used as template. In 'ΔHalfvol1–4' ΔHalfvol1–4 cells taken from a plate were used as template, while '-' is the negative control. (B) The position of the primers employed in the PCR are indicated by red arrows. A PCR product is generated exclusively when the viral DNA has circularised. (C) Sanger sequencing of the bands obtained in the PCR from the 'Supernatant' samples. The *attP* sequence is highlighted in light blue.

Fig. S8, strain Δ*oapA*), and (b) mass spectrometry to detect viral proteins (Table S5).

In order to demonstrate provirus excision and release from the cell, isolation of viruses from the culture supernatant was performed. Proviruses Halfvol1, Halfvol2, and Halfvol3 are flanked by 14 nt direct repeats, and circularize upon excision. Virus circularization was validated using PCR with outward-facing primers positioned close to the provirus ends (Fig. 5). PCR products that traverse the *attP* site were confirmed by Sanger sequencing. Intact *attP* sites indicate a circularised virus genome, thus

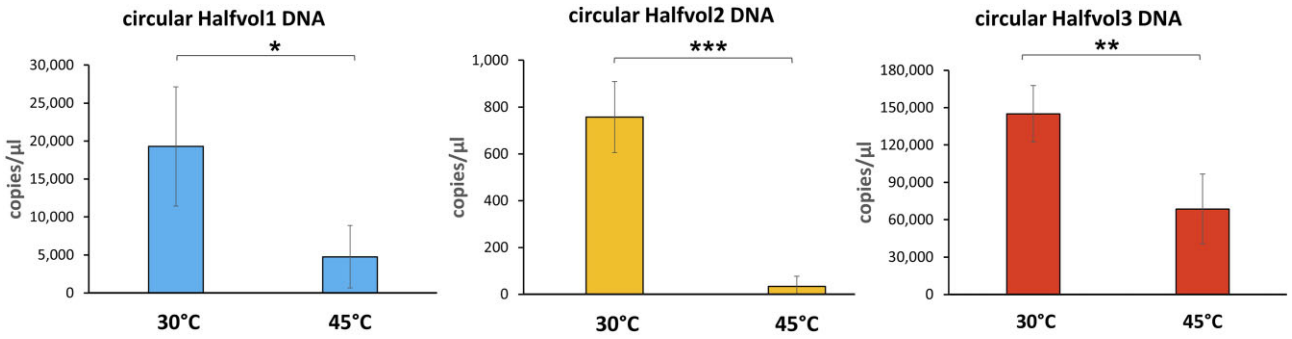

Figure 6. Quantification of circular Halfvol1, Halfvol2, and Halfvol3 DNA at different growth temperatures. qPCR was employed to determine the concentration of circular Halfvol1, Halfvol2, and Halfvol3 DNA in the culture supernatant of the provirus-containing strain Δ*oapA*, cultured at 30°C (low growth temperature for *H. volcanii*) or 45°C (standard growth temperature) during the stationary phase (OD_{650} 0.86). Results have been taken from three biological replicates. Asterisks indicate significant differences (*t*-test); ∗: $P \leq .05$; ∗∗: $P \leq .01$; and ∗∗∗: $P \leq .0001$.

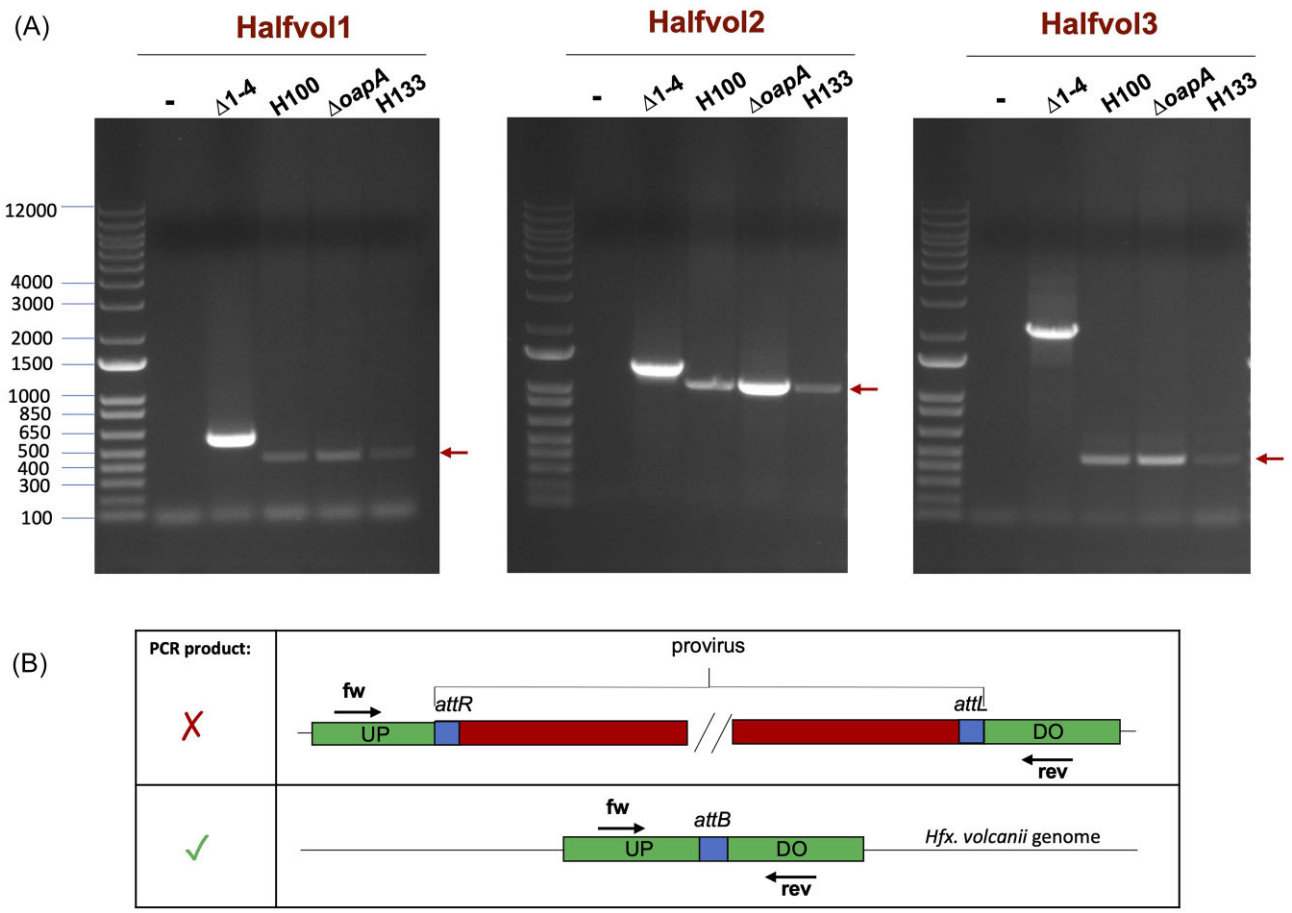

Figure 7. Detection of *attB* sites, indicating provirus excision. (A) The excision of Halfvol1, Halfvol2, and Halfvol3 from the chromosome is shown in different *H. volcanii* strains (Δ1–4, H100, Δ*oapA*, and H133) by PCR. For Halfvol1, spontaneous excision results in a 451-bp signal (weak signals in lanes H100, Δ*oapA*, and H133, indicated by an arrow) whereas the deletion strain (ΔHalfvol1–4: Δ1–4) shows a band at 585 bp. For Halfvol2, spontaneous excision gives a signal of 1015 bp (indicated by an arrow), whereas the deletion strain (Δ1–4) gives a signal of 1263 bp. For Halfvol3, spontaneous excision results in a 465-bp signal (indicated by an arrow), while the deletion strain (Δ1–4) shows a band at 2172 bp. The size increase is due to insertion of the *trpA* marker. (B) The primers employed in the PCR are indicated by black arrows and they are located in the upstream (UP) and downstream (DO) regions of the specific provirus. A PCR product is generated exclusively when the provirus excised from the *H. volcanii* chromosome.

confirming virus excision. Interestingly, provirus excision was more efficient when the culture was grown at low temperature (30°C) (Fig. 6). The qPCR titre of Halfvol3 (see y-axis in Fig. 6) is higher than for Halvol1 and Halfvol2 suggesting a more efficient excision of Halfvol3. We did not detect any circularization of Halfvol4.

Given the presence of circular virus DNA, the occurrence of provirus-free *attB* sites was also examined via PCR in three

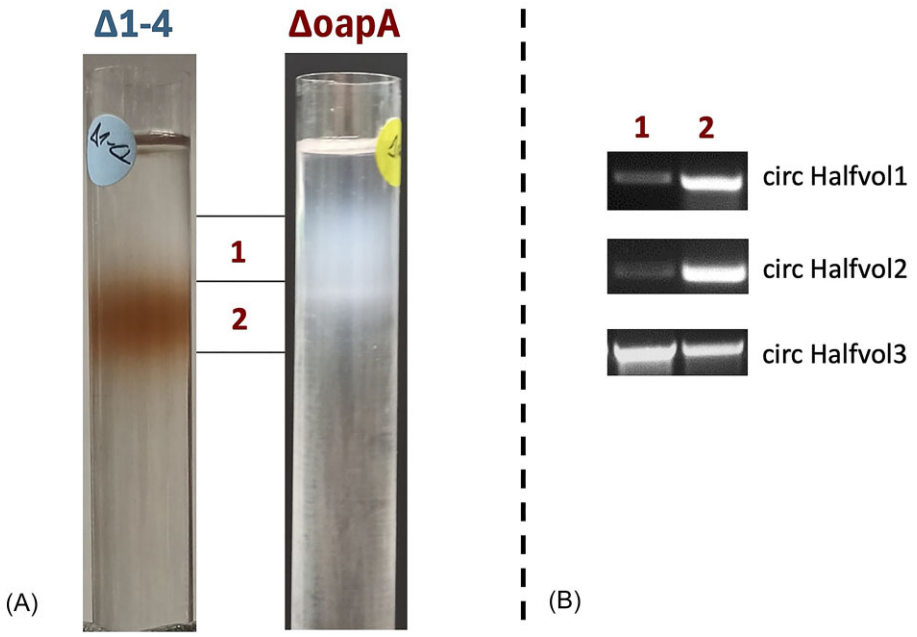

Figure 8. Enrichment of isolated excised provirus from **Δ***oapA* supernatant by CsCl density-gradient centrifugation. (A) CsCl density-gradient ultracentrifugation for samples obtained from strains ΔHalfvol1–4 (Δ1–4, lacking proviruses) and Δ*oapA* (containing proviruses). The ΔHalfvol1–4 sample (left tube) shows a broad pink band while the Δ*oapA* sample reveals one broad white band and one distinct white band (right tube). (B) PCR confirms the presence of circularised viral DNA in the first and second bands ('1' and '2') of the Δ*oapA* sample. Circular forms of Halfvol1, Halfvol2, and Halfvol3 DNA were found in fractions of both bands.

provirus-containing strains, which were used in this study: the control strain H100, the parent strain H133, and in addition strain Δ*oapA*, which has a deletion in gene *oapA* (HVO_3014) and produces only a small amount of extracellular vesicles (EVs) (see below) (Mills et al. 2024). Convergent primers targeting the upstream and downstream region of the provirus were used. Given the substantial size of the proviral regions (12–20 kb), a PCR product would only be expected if the provirus had been naturally excised from the chromosome (Fig. 7). PCR products of the anticipated size were obtained and subjected to Sanger sequencing, which confirmed the existence of provirus-free *attB*, and so validated that the proviruses had been excised from the chromosome.

Identification of potential virus particles

Virus particles can be enriched by subjecting the culture supernatant to CsCl density-gradient centrifugation. Due to the production of EVs, *H. volcanii* naturally shows a strong pink band in such gradients (Mills et al. 2024). Since EVs interfere with virus detection, we used strain Δ*oapA*, which has a largely reduced EV production by deletion of the gene coding for GTPase HVO_3014 (*oapA*; also called *arvA*) (Mills et al. 2024). This strain was used to enrich virus particles upon CsCl density-gradient centrifugation and to investigate the gradient fraction for the presence of virus particles using electron microscopy. After CsCl density-gradient centrifugation of the virus preparation from Δ*oapA*, two distinct whitish bands were visible (Fig. 8A, right tube). Whereas the ΔHalfvol1–4 preparation, which produces EVs, showed a singular wide pink band (Fig. 8A, left tube). The presence of viral DNA (circular Halfvol1, Halfvol2, and Halfvol3) was confirmed by PCR in both bands of the Δ*oapA* sample (Fig. 8B).

Samples from the two bands in Δ*oapA*, and the corresponding parts in ΔHalfvol1–4 as control, were subjected to TEM analysis. Within the first band, mainly two different particles were observed (Fig. 9A).

Unstructured round particles (diameter of ~40 nm) were observed in Δ*oapA* and ΔHalfvol1–4, which probably represent EVs. In Δ*oapA*, larger and more structured particles (diameter of ~80 nm) were detected. These latter particles might represent virus particles originating from activated proviruses. A set of representative images is shown in Fig. S8. In the second band of the Δ*oapA* sample, spherical particles attached to tail-like structures were observed (Fig. 9B). The spherical particles were uniform in size, characterised by a bright white core surrounded by a darker halo. The tail-like entities appeared to be structured, with a notable variability in length, with the longest measuring up to ~1.7 µm. Since none of the provirus regions contain genetic evidence for tailed viruses, it remains to be determined if the tail-like structures are virus tails or if they are cellular structures (archaella or pili) to which virus particles have attached. A series of representative images is shown in Fig. S8.

Analysis of density-gradient centrifugation bands by mass spectroscopy

The second band from CsCl density-gradient centrifugation derived from strain Δ*oapA* and its counterpart from strain ΔHalfvol1–4 were loaded on a 12% SDS gel and were subjected to mass spectrometry analysis. Viral proteins were identified in the Δ*oapA* sample (Table 5, Table S5) but not in the ΔHalfvol1–4 sample, that was analysed as a control. Four different proteins associated with Halfvol1 were identified, one of them being HVO_0271, the homolog of structural protein VP4. Seven different proteins associated with Halfvol3 were identified, one of them being a VP4 homologue (HVO_1431) with a total spectrum count of 69 (Table 5). Quantification of the excised circularised viruses with qPCR (Fig. 6) showed the highest titre for Halfvol3 suggesting that this provirus is excised more efficiently. The mass spectrometry data also show higher total spectrum counts for Halfvol3 derived proteins, supporting this hypothesis. Remarkably, nineteen different

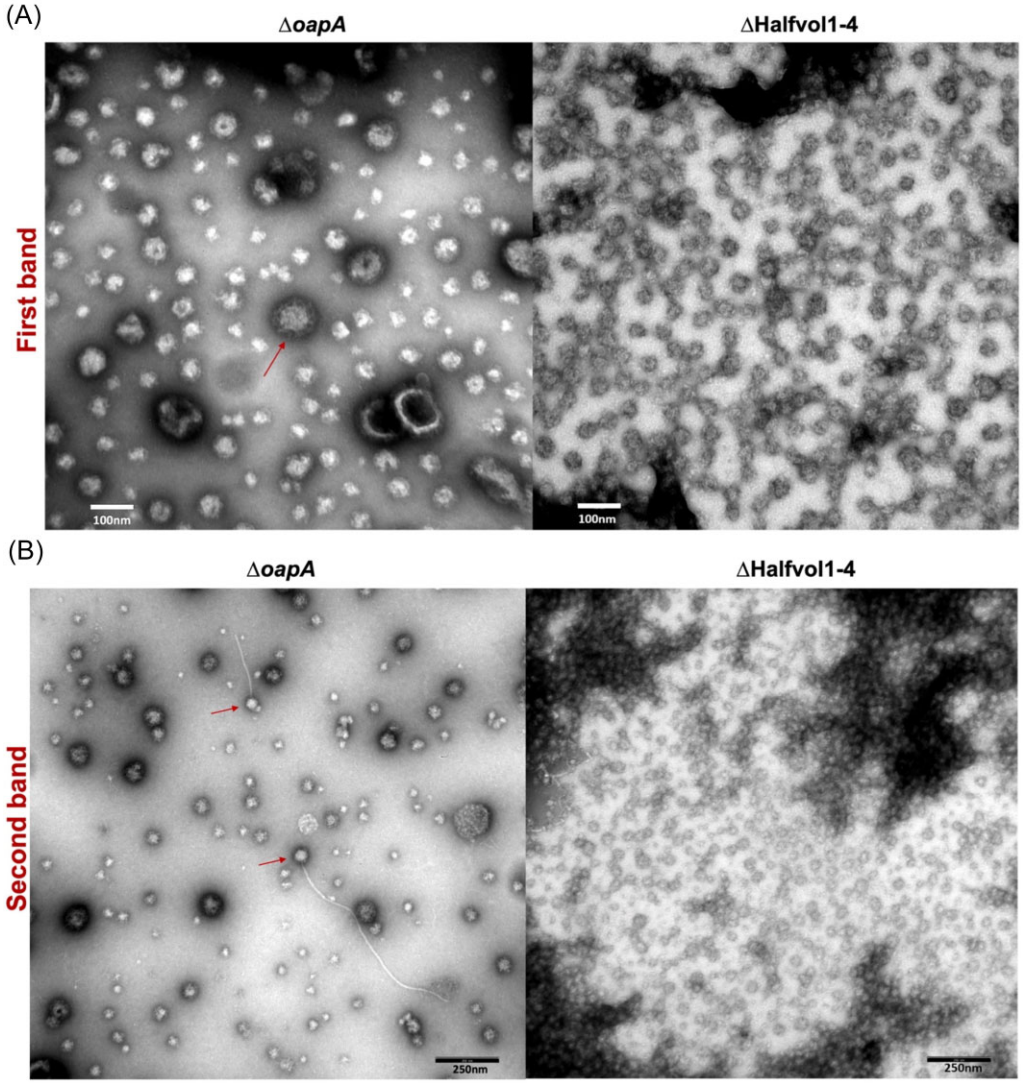

Figure 9. TEM images of purified supernatant from provirus-containing strain Δ*oapA* and provirus-free strain ΔHalfvol1–4. The isolation of virus sample was performed in the provirus-containing strain (Δ*oapA*) and the same procedure was carried out in the provirus-free strain (ΔHalfvol1–4) as a control. (A) First band of the CsCl density gradient from strain Δ*oapA* and its counterpart from strain ΔHalfvol1–4. A structure that resembles a pleolipovirus, complete with spike proteins, is highlighted with a red arrow. The scale bar indicates 100 nm. (B) Second band of the CsCl density gradient from strain Δ*oapA* and its counterpart from strain ΔHalfvol1–4. Head-tailed structures were observed in the Δ*oapA* sample (red arrows). The scale bar indicates 250 nm.

proteins were identified for Halfvol4, featuring the highest number of total spectrum counts compared to the other viruses.

But as stated below, none of the Halfvol4 proteins showed similarity to known viral proteins. However, provirus proteins (from all four proviruses) were not the most abundant proteins identified in the CsCl fraction. These were membrane-associated proteins such as ABC-transporters, S-layer protein, halocyanin, oxidoreductases, FtsZ, and PilA (Table S5). A similar observation was made upon analysis of EVs of *H. volcanii*; isolation of EVs and of the pleolipovirus HFPV-1 in *H. volcanii* and subsequent identification of proteins in these fractions revealed the presence of many membrane-associated proteins (Alarcon-Schumacher et al. 2022, Mills et al. 2024). Since our fractions were obtained from a strain with reduced EV production (Δ*oapA*) these cellular proteins very likely are derived from cell debris present in the culture supernatant.

Attempts to infect the ΔHalfvol1–4 strain with isolated viral particles were conducted using both liquid culture infection and plaque assay (data not shown). However, these experiments were unsuccessful. This may be due to the limited abundance of proviruses, as shown by the mass spectrometry analyses, which revealed low abundance of viral proteins in virus preparations.

Discussion

Four putative provirus regions have been identified in the *H. volcanii* genome (Halfvol1, Halfvol2, Halfvol3, and Halfvol4), which have not been extensively studied previously. Our research focused on characterizing these proviruses, investigating their potential activity, and understanding their impact on *H. volcanii*.

The provirus deletion strain is phenotypically different from a provirus-containing strain

Since we could delete all four provirus regions, none of them has evolved to an essential part of the genome under the growth

Table 5. Provirus proteins identified in purified fractions. Fraction 2 of the CsCl gradient from Δ*oapA* and ΔHalfvol1–4 cells were analysed by mass spectrometry. Identified provirus proteins are listed, for which the total spectrum count was higher in Δ*oapA* samples compared to Δ1–4 samples. An average of the total spectrum count values from the three biological replicates was calculated. Column Δ*oapA*: average total spectrum counts detected in the Δ*oapA* sample, column ΔHalfvol1–4: average total spectrum counts detected in the ΔHalfvol1–4 sample. The complete results of this analysis are shown in Table S5.

Gene	Function	Δ*oapA*	ΔHalfvol1–4
Halvol1			
HVO_0277	Conserved hypothetical protein	6	0
HVO_0272	Homolog to HFPV1-VP5	5	0
HVO_0271	**Structural protein VP4 (spike protein)**	**3**	**0**
HVO_0259	Conserved hypothetical protein	2	0
Halfvol2			
HVO_0379A	Conserved hypothetical protein	19	0
HVO_0379B	Conserved hypothetical protein	19	0
Halfvol3			
HVO_1431	**Structural protein VP4 (spike protein)**	**69**	**0**
HVO_1424	Conserved hypothetical protein	16	0
HVO_1425	Conserved hypothetical protein	10	0
HVO_1429	Homolog to HRPV2-VP6	5	0
HVO_1428	VP8-like ATPase	3	0
HVO_1433	Homolog to pHK2-ORF2	3	0
HVO_1426	VP8-like ATPase	2	0
Halfvol4			
HVO_2286	Conserved hypothetical protein	208	2
HVO_2263	HTH domain protein	117	0
HVO_2269	RmeR	51	1
HVO_2267	Conserved hypothetical protein	29	0
HVO_2266	Conserved hypothetical protein	22	0
HVO_2256	DUF262 family protein	17	0
HVO_2271	RmeS	16	0
HVO_2279	Conserved hypothetical protein	15	0
HVO_2270	RmeM	14	0
HVO_2276	Avs2 family antiviral STAND domain protein	14	0
HVO_2265	Conserved hypothetical protein	11	0
HVO_2253	Conserved hypothetical protein	10	0
HVO_2288	Conserved hypothetical protein	8	0
HVO_2291	Conserved hypothetical protein	6	0
HVO_2255	Conserved hypothetical protein	5	0
HVO_2275	N-6 adenine-specific DNA methylase domain protein	4	3
HVO_2284A	Hypothetical protein	2	0
HVO_2254	Conserved hypothetical protein	1	0
HVO_2268	TATA-binding transcription initiation factor	1	0

conditions used in the laboratory. A phenotypic comparison between the strain containing proviruses (H100) and the provirus-free strain (ΔHalfvol1–4) revealed various differences across several physiological parameters. ΔHalfvol1–4 showed a slightly faster growth compared to the control (Fig. 1 and Fig. S4) under the majority of the conditions tested (low salt, low and high temperature). We propose that the production of viral proteins and particles in H100 contributes to its reduced growth compared to the provirus-free strain. Also, the production of viral particles could interfere with membrane processes, such as the electron transport chain or transport of substrates within the cell. This may account for the observed variations in membrane-associated protein expression in the ΔHalfvol1–4 strain, as highlighted by the DAVID analysis of the transcriptome data. Numerous studies have documented the impact of nonlytic viral infections on strain growth, often associated with growth retardation. A recent study identified a natural *Haloferax* strain (48 N) closely related to the model organism *H. volcanii* (Turgeman-Grott et al. 2024). This strain was chronically infected by a lemon-shaped virus. Deleting this virus from 48 N resulted in significant alterations in its gene expression profile and a substantial increase of its growth rate (Turgeman-Grott et al. 2024). Another well-known example

is the nonlytic bacteriophage M13, which infects *Escherichia coli* and reduces its growth rate by 15% according to Kick et al. (2017) or to 60% of the growth rate of uninfected cells, as reported by Wan and Goddard (2012).

In addition, ΔHalfvol1-4 exhibits higher resistance to oxidative stress in comparison to the control (Fig. 4), while the UV stress test showed no difference between the deletion strain and control strain. Hydrogen peroxide (H_2O_2) is a reactive oxygen species (ROS) that can cause significant damage to cells, including DNA, proteins, and lipids, leading to broad-ranging damage. In contrast, UV radiation primarily induces damage to DNA. Therefore, provirus genes seem to interfere with resistance against ROS. Virus infection is known to induce oxidative stress in the infected cells (Su et al. 2013, Ivanov et al. 2017). ROS response might include increased activity of virus defence systems, therefore it would be beneficial for a virus to block or reduce oxidative stress response. The genes which are responsible for the observed effects will be subject of future analyses.

Furthermore, ΔHalfvol1–4 cells display an elongated shape (Fig. 2), a characteristic often associated with motility in bacteria and archaea (Schwarzer et al. 2021, Schiller et al. 2024) suggesting that the proviruses have an influence on the host motility machinery. A correlation between the motility and oxidative stress response has been discovered during interaction between *Campylobacter jejuni* and its virus NCTC 12673 (Sacher et al. 2021).

Proviruses affect motility and CRISPR-Cas expression providing thereby a potential superinfection exclusion mechanism

RNA-seq analysis showed upregulation of archaella, chemotaxis, and pili genes in ΔHalfvol1–4, which are all involved in motility (Table 3 and Table S4). This result provides a molecular explanation for the hypermotility of ΔHalfvol1–4 observed in the swarming assay (Fig. 3). Numerous instances in the literature highlight the influence of viruses or proviruses on cellular motility behaviour and the expression of archaella. For instance, transcriptome analysis of *H. volcanii* infected with the pleolipovirus HFPV-1 revealed an overall downregulation of archaella genes (Alarcon-Schumacher et al. 2022). Similarly, the deletion of the prophage CP4-57 from *E. coli* induced motility through flagella expression, as evidenced by transcriptome analysis and swarming assay results (Wang et al. 2009). Additionally, Brathwaite and colleagues reported impaired flagellar motility in *Campylobacter* due to the presence of the phage carrier state (Brathwaite et al. 2015).

Proviruses also seem to affect the immune system of *H. volcanii*. RNA-seq analysis reveals a significant downregulation of CRISPR arrays (-3 or -2 log2FC) in the strain lacking proviruses, suggesting an interplay between proviruses and the CRISPR-Cas virus defence system (Table 3). Interestingly, Zegans et al. (2009) observed that a functional CRISPR-Cas system is necessary to inhibit swarming motility of *P. aeruginosa* by the DMS3 prophage (Zegans et al. 2009).

Proviruses frequently encode superinfection exclusion mechanisms. We propose that the proviruses present in *H. volcanii* likely confer protection against other viruses through such mechanisms. Our hypothesis is based on the observation that the wild-type strain containing proviruses exhibits higher expression of crRNAs compared to the provirus-free strain, as revealed by RNA sequencing data. This indicates that the presence of proviruses enhances the immune response by increasing crRNA production, thereby protecting *H. volcanii* against other invading viruses. Furthermore, it appears that the presence of proviruses suppresses the expression of archaella/pili genes in *H. volcanii*. This inhibition could prevent other viruses from entering the host cells, if they rely on these structures for attachment and adsorption. Numerous viruses utilize flagella or pili for cellular entry (Tittes et al. 2021b). For instance, bacteriophage x has been observed to attach to the flagella of *E. coli* via a tail fibre, subsequently migrating to the base of the flagella to inject its DNA (Schade et al. 1967). Similarly, ΦCB13 and ΦCbK viruses, infecting *Caulobacter crescentus*, adhere to the host's flagella through a filament on the phage head (Guerrero-Ferreira et al. 2011), while phage 7–7-1, infecting *Agrobacterium sp.* strain H13-3, utilizes tail fibres to attach to the flagella for absorption (Lotz et al. 1977).

In this context, proviruses within *H. volcanii* may downregulate archaella expression as a mechanism to prevent other viruses from entering and infecting the host. This dynamic hints at a potential mutualistic relationship, wherein lysogeny benefits the chronic virus by preventing further infections from competing viruses, while simultaneously increasing the host protection against more virulent viruses through superinfection exclusion mechanisms. Moreover, the observed reduction in host motility, often perceived as negative, could serve as a strategy for self-quarantine: by restricting its own movement, *H. volcanii* minimizes virus spread and thereby reducing the chance of virus transmission. This may be a general strategy employed by hosts infected by a persistent virus (Chung et al. 2014).

Additionally, the Halfvol4 region contains several defence systems: a restriction modification system (rmeR, rmeM, and rmeS) and the Avast2 protein (Fig. S1), thus enhancing the protection of *H. volcanii* against other invading viruses.

The superinfection exclusion effect of the provirus Halfvol3 against the pleolipovirus HFPV-1 has been documented (Alarcon-Schumacher et al. 2022). It was shown that the strain ΔHalfvol3 is more susceptible to HFPV-1 compared to the control. This suggests that the provirus region likely harbours one or more genes involved in mitigating the impact of HFPV-1. Moreover, in this study, the infection of HFPV-1 in a wild-type *H. volcanii* strain led to a notable downregulation of genes of Halfvol1. Specifically, VP3 was suppressed over 150-fold compared to uninfected cultures, while VP4, HVO_0268, and HVO_0270 exhibited decreases in expression with fold changes ranging between 40 and 60. These findings suggest that Halfvol1 may trigger a defensive response against HFPV-1, highlighting the intricate interplay of virus–virus interactions within the host (Alarcon-Schumacher et al. 2022).

Chronic infection is widespread within the archaeal domain. It has been suggested that in archaea, chronic infection serves as a strategic trade-off in which the host cell maintains a persistent viral infection to protect itself from a more dangerous virus that could otherwise kill it (Wirth and Young 2020, Turgeman-Grott et al. 2024).

Given the rarity of *H. volcanii* viruses and the knowledge that proviruses often carry genes for superinfection exclusion mechanisms aimed at safeguarding the host from other viral threats, it is conceivable that the four proviruses present within *H. volcanii* confer some level of superinfection exclusion. Consequently, the provirus-free strain may serve as a more susceptible host to viral infections, a factor with potential implications for future virology research in *H. volcanii*.

Viruses are generated by provirus circularization and excision

Our results confirm previous data that Halfvol1 and Halfvol2 can be activated in *H. volcanii* (Dyall-Smith et al. 2021). Excision of viral DNA and its presence in the supernatant were confirmed through PCR and Sanger sequencing. Specifically, circular DNA of Halfvol1, Halfvol2, and Halfvol3 (*attP* site) was detected in the culture supernatant (Fig. 5). Further evidence of provirus excision is provided by generation of PCR products across the chromosomal *attB* site, confirmed by sequencing (Fig. 7). This demonstrates that *H. volcanii* strains possess chromosomal copies lacking proviruses and in addition confirms the spontaneous excision of proviruses from the chromosome, suggesting a persistent infection mode for these viruses.

Provirus activation was predominantly observed when cultures are grown at 30°C, suggesting that these viruses were more active at temperatures below the host's optimal growth temperature (45°C) (Fig. 6). There are several documented instances, where viruses exhibit increased activity at lower temperatures. For example, the archaeal pleomorphic virus AvPV1 shows higher titres at lower temperatures (Baquero et al. 2024), and the pleolipovirus HFPV1 was isolated from *H. volcanii* culture grown at 28°C (Alarcon-Schumacher et al. 2022). It was demonstrated that elevated temperatures can impact phage infectivity by altering the structural integrity and flexibility of their lipid membranes or capsid proteins, which may influence their lifecycle strategies (Zhang et al. 2022). Temperature shifts are known to potentially trigger viral activity (Schuster et al. 1973, Choi et al. 2010, Makky et al. 2021, Zhang et al. 2022), indicating a possible ecological interaction between host and virus during colder periods. Numerous studies suggest that phage life cycle strategies may vary with the seasons. For instance, research in Antarctic Salt Lake and Tampa Bay has reported a higher prevalence of lysogeny during winter and spring compared to summer (Zhang et al. 2022). It is important to note that seasonal variations involve changes in several environmental factors, such as temperature, salinity, and primary productivity, highlighting the complex interplay of elements that influence phage lifecycle dynamics on a seasonal basis (Zhang et al. 2022).

Electron microscopy analysis of virus enriched fractions from the EV-reduced strain revealed that the first band contained, besides residual EVs, structured particles which were quite large (~80 nm in diameter) and were not observed in ΔHalfvol1–4 (Fig. 9A). A collection of such particles illustrates that they are quite uniform and thus are candidates to represent pleolipovirus structures (Fig. S8). The second band contained spherical particles with a long tail, which were not observed in ΔHalfvol1–4 (Fig. 9B, Fig. S8). We cannot link genes from any of the proviruses to these structures. Therefore, it is not possible to say from which provirus they originate. It is yet unresolved if the 'tails' represent virus tails or if they represent cellular surface structures to which virus particles have attached. Thus, the particles isolated from the supernatants of Δ*oapA* cultures might be activated proviruses of *H. volcanii*. This hypothesis is supported by the considerable differences observed between the Δ*oapA* and ΔHalfvol1-4 samples by TEM. While particles being small and heterogenous are found in both samples, the larger and more uniform particles were observed exclusively in the Δ*oapA* strain. Additionally, mass spectrometry analysis of the CsCl density fractions of Δ*oapA* revealed the presence of viral proteins (Table 5). However, since we did not succeed in infecting ΔHalfvol1–4 with isolated virus particles we cannot state at this point that we have active viral particles. The head–tailed structure observed in the Δ*oapA* supernatant (Fig. 9B.) could be viral particles (pleolipovirus encoded by Halfvol1 or Halfvol3, or the novel group virus encoded by Halfvol2) associating with fragments of archaella/pili or some other type of filament from *H. volcanii*. According to this hypothesis, the virus would bind to archaella or pili during the adsorption step to facilitate entry into the host cell.

Taken together, the results indicate that *H. volcanii* proviruses excise from the genome and replicate at low level, causing a type of persistent infection.

Is Halfvol4 a provirus or an integrated plasmid?

Halfvol4 was initially described as a provirus (Norais et al. 2007, Hartman et al. 2010). However, upon closer examination of the Halfvol4 genome, it was not possible to detect any candidates for capsid proteins. Also, none of the proteins showed significant similarity to viral proteins. Halfvol4 exhibits characteristics of integrative and conjugative elements (ICEs). Interestingly, Halfvol4 encodes a T4SS, which functions similarly to plasmid conjugation systems, facilitating the transfer of ICEs between cells (Johnson and Grossman 2015). ICEs range in size from 18 to 500 kb and are typically AT-rich compared to their host genome. They often carry cargo genes that provide phenotypic traits to host cells, such as antibiotic resistance, pathogenesis, restriction modification (as also seen in Halfvol4), and carbon-source utilization (Johnson and Grossman 2015). It is important to note that ICEs are usually flanked by *att* sites and form circular DNA upon excision. A slightly extended Halfvol4 (as compared to its original description) is enclosed in a 14-nt direct repeat, which resembles an *att* site. An additional copy of the 14-nt sequence is encountered internally within Halfvol4. Halfvol4 appears to be inserted into the *H. volcanii* genome sequence when compared to some closely related haloarchaeal strains that do not have an insertion at this position. A few other strains have inserts at the corresponding position however these inserts have a different sequence. In *Hfx. alexandrinus* pws11, regions within the insert show similarity to Halfvol4, in *Haloferax* sp. Atlit-105R the insert is totally unrelated (Table S2). Circular Halfvol4 DNA was not detected in the cell fraction or culture supernatant. Therefore, we suggest that Halfvol4 is not a provirus but a plasmid that has been integrated into the *H. volcanii* chromosome.

Acknowledgements

We thank Manuela Weishaupt, Susanne Schmidt, Elena Katzowitsch, and Panagiota Arampatzi for expert technical assistance as well as Jessika Jakubowski and Sandra Schreiber for help with generating the provirus deletion strain ΔHalfvol3ΔHalfvol4. For help with uploading proteome and RNA-seq data to the respective repositories we thank Olexandr Dybkov and Lisa-Katharina Maier. In addition, we thank Lisa-Katharina Maier and Uri Gophna for constructive and fruitful discussions as well as for careful reading of the manuscript.

Author contributions

N.D.C., S.B., J.B., M.R., S.K., T.A., and L.M. performed experiments; N.D.C., S.B., J.B., S.K., T.B., T.G., C.R., P.W., T.A., H.U., S.E., M.D.S., F.P., and A.M. analysed the samples and the data. N.D.C. and A.M. conceptualised the project and A.M. supervised the research. H.U., T.A., and A.M. provided resources and funding. N.D.C., F.P., M.D.S.,

and A.M. wrote the original draft. All authors reviewed and edited the draft and approved the submitted version.

Supplementary data

Supplementary data are available at *FEMSML* online.

Conflict of interest: The authors do not have any conflict of interests.

Funding

Work in the laboratory of Anita Marchfelder was funded by the DFG (Ma1538/25-2, DFG priority programme "CRISPR-Cas functions beyond defence" SPP2141 and Ma1538/27-1). Work in the laboratory of Thorsten Allers was funded by The Leverhulme Trust (RF-2023-286\2). The Core Unit Systems Medicine is partly funded (Z-6) by the Interdisciplinary Center for Clinical Research (IZKF) Würzburg.

Data availability

Transcriptome data was deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under study accession number: PRJEB83039. Proteome data was deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Vizcaíno et al. 2016) with the dataset identifier: PXD058393.

References

Alarcon-Schumacher T, Naor A, Gophna U *et al.* 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.

Allers T, Ngo HP, Mevarech M *et al.* Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. *Appl Environ Microbiol* 2004;**70**:943–53.

Atanasova NS, Demina TA, Krishnam Rajan Shanthi SNV *et al.* Extremely halophilic pleomorphic archaeal virus HRPV9 extends the diversity of pleolipoviruses with integrases. *Res Microbiol* 2018;**169**:500–4.

Atanasova NS, Oksanen HM, Bamford DH. Haloviruses of archaea, bacteria, and eukaryotes. *Curr Opin Microbiol* 2015;**25**:40–48.

Atanasova NS, Senčilo A, Pietilä MK *et al.* Comparison of lipid-containing bacterial and archaeal viruses. *Adv Virus Res* 2015b;**92**:1–61.

Ausiannikava D, Mitchell L, Marriott H *et al.* Evolution of genome architecture in Archaea: spontaneous generation of a new chromosome in *Haloferax volcanii*. *Mol Biol Evol* 2018;**35**:1855–68.

Badel C, Da Cunha V, Oberto J. Archaeal tyrosine recombinases. *FEMS Microbiol Rev* 2021;**45**:fuab004.

Baquero DP, Bignon EA, Krupovic M. Pleomorphic viruses establish stable relationship with marine hyperthermophilic archaea. *ISME J* 2024;**18**:wrae008.

Beverley SM. Estimation of circular DNA size using gamma-irradiation and pulsed-field gel electrophoresis. *Anal Biochem* 1989;**177**:110–4.

Bitan-Banin G, Ortenberg R, Mevarech M. Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the pyrE gene. *J Bacteriol* 2003;**185**:772–8.

Bondy-Denomy J, Davidson AR. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. *J Microbiol* 2014;**52**:235–42.

Brathwaite KJ, Siringan P, Connerton PL *et al.* Host adaption to the bacteriophage carrier state of *Campylobacter jejuni*. *Res Microbiol* 2015;**166**:504–15.

Campbell AM. Chromosomal insertion sites for phages and plasmids. *J Bacteriol* 1992;**174**:7495–9.

Casjens S. Prophages and bacterial genomics: what have we learned so far?. *Mol Microbiol* 2003;**49**:277–300.

Choi J, Kotay SM, Goel R. Various physico-chemical stress factors cause prophage induction in *Nitrosospira multiformis* 25196- an ammonia oxidizing bacteria. *Water Res* 2010;**44**:4550–8.

Chung I-Y, Jang H-J, Bae H-W *et al.* A phage protein that inhibits the bacterial ATPase required for type IV pilus assembly. *Proc Natl Acad Sci USA* 2014;**111**:11503–8.

de Silva RT, Abdul-Halim MF, Pittrich DA *et al.* Improved growth and morphological plasticity of Haloferax volcanii. *Microbiology* 2021;**167**:001012. https://doi.org:https://doi.org/10.1099/mic.0.001012

Delmas S, Shunburne L, Ngo HP *et al.* Mre11-Rad50 promotes rapid repair of DNA damage in the polyploid archaeon *Haloferax volcanii* by restraining homologous recombination. *PLoS Genet* 2009;**5**:e1000552.

Demina TA, Oksanen HM. Pleomorphic archaeal viruses: the family *Pleolipoviridae* is expanding by seven new species. *Arch Virol* 2020;**165**:2723–31.

Dyall-Smith M, Pfeiffer F, Chiang PW *et al.* The novel halovirus Hardycor1, and the presence of active (induced) proviruses in four haloarchaea. *Genes* 2021;**12**:149.

Dyall-Smith ML. *The Halohandbook: Protocols for Halobacterial Genetics*. 2009.

Feiner R, Argov T, Rabinovich L *et al.* A new perspective on lysogeny: prophages as active regulatory switches of bacteria. *Nat Rev Microbiol* 2015;**13**:641–50.

Folimonova SY. Superinfection exclusion is an active virus-controlled function that requires a specific viral protein. *J Virol* 2012;**86**:5554–61.

Förstner KU, Vogel J, Sharma CM. READemption-a tool for the computational analysis of deep-sequencing-based transcriptome data. *Bioinformatics* 2014;**30**:3421–3.

Gao L, Altae-Tran H, Böhning F *et al.* Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* 2020;**369**:1077–84.

Guerrero-Ferreira RC, Viollier PH, Ely B *et al.* Alternative mechanism for bacteriophage adsorption to the motile bacterium *Caulobacter crescentus*. *Proc Natl Acad Sci USA* 2011;**108**:9963–8.

Hadjeras L, Bartel J, Maier LK *et al.* Revealing the small proteome of *Haloferax volcanii* by combining ribosome profiling and small-protein optimized mass spectrometry. *Microlife* 2023;**4**:uqad001.

Harshitha R, Arunraj DR. Real-time quantitative PCR: a tool for absolute and relative quantification. *Biochem Mol Bio Edu* 2021;**49**:800–12.

Hartman AL, Norais C, Badger JH *et al.* The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. *PLoS One* 2010;**5**:e9605.

Hawkins M, Malla S, Blythe MJ *et al.* Accelerated growth in the absence of DNA replication origins. *Nature* 2013;**503**:544–7.

Hoffmann S, Otto C, Kurtz S *et al.* Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput Biol* 2009;**5**:e1000502.

Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;**4**:44–57.

Hunter M, Fusco D. Superinfection exclusion: a viral strategy with short-term benefits and long-term drawbacks. *PLoS Comput Biol* 2022;**18**:e1010125.

Ivanov AV, Bartosch B, Isaguliants MG. Oxidative stress in infection and consequent disease. *Oxid Med Cell Long* 2017;**2017**:3496043.

Johnson CM, Grossman AD. Integrative and conjugative elements (ICEs): what they do and how they work. *Annu Rev Genet* 2015;**49**:577–601.

Ken R, Hackett NR. *Halobacterium halobium* strains lysogenic for phage phi H contain a protein resembling coliphage repressors. *J Bacteriol* 1991;**173**:955–60.

Kick B, Behler KL, Severin TS *et al.* Chemostat studies of bacteriophage M13 infected *Escherichia coli* JM109 for continuous ssDNA production. *J Biotechnol* 2017;**258**:92–100.

Laass S, Monzon VA, Kliemt J *et al.* Characterization of the transcriptome of *Haloferax volcanii*, grown under four different conditions, with mixed RNA-seq. *PLoS One* 2019;**14**:e0215986.

Leigh JA, Albers SV, Atomi H *et al.* Model organisms for genetics in the domain Archaea: methanogens, halophiles, thermococcales and sulfolobales. *FEMS Microbiol Rev* 2011;**35**:577–608.

Liu Y, Wang J, Liu Y *et al.* Identification and characterization of SNJ2, the first temperate pleolipovirus integrating into the genome of the SNJ1-lysogenic archaeal strain. *Mol Microbiol* 2015;**98**:1002–20.

Lotz W, Acker G, Schmitt R. Bacteriophage 7-7-1 adsorbs to the complex flagella of *Rhizobium lupini* H13-3. *J Gen Virol* 1977;**34**:9–17.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;**15**:550.

Luk AW, Williams TJ, Erdmann S, Papke RT, Cavicchioli R(2014) Viruses of haloarchaea. Life (Basel) 4: 681-715. https://doi.org:10.3390/life4040681

Maier LK, Stachler AE, Brendel J *et al.* The nuts and bolts of the Haloferax CRISPR-Cas system I-B. *RNA Biol* 2019;**16**:469–80.

Makky S, Dawoud A, Safwat A *et al.* The bacteriophage decides own tracks: when they are with or against the bacteria. *Curr Res Microb Sci* 2021;**2**:100050.

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 2011;**17**:3.

McKerral JC, Papudeshi B, Inglis LK *et al.* The promise and pitfalls of prophages. *bioRxiv* 2023. http://dx.doi.org/10.1101/2023.04.20.537752

Mills J, Gebhard LJ, Schubotz F *et al.* Extracellular vesicle formation in Euryarchaeota is driven by a small GTPase. *Proc Natl Acad Sci USA* 2024;**121**:e2311321121.

Mizuno CM, Prajapati B, Lucas-Staat S *et al.* Novel haloarchaeal viruses from Lake Retba infecting *Haloferax* and *Halorubrum* species. *Environ Microbiol* 2019;**21**:2129–47.

Munson-McGee JH, Peng S, Dewerff S *et al.* A virus or more in (nearly) every cell: ubiquitous networks of virus–host interactions in extreme environments. *ISME J* 2018;**12**:1706–14.

Norais C, Hawkins M, Hartman AL *et al.* Genetic and physical mapping of DNA replication origins in *Haloferax volcanii*. *PLoS Genet* 2007;**3**:e77.

Nuttall SD, Dyall-Smith ML. HF1 and HF2: novel bacteriophages of halophilic archaea. *Virology* 1993;**197**:678–84.

Payne LJ, Meaden S, Mestre MR *et al.* PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. *Nucleic Acids Res* 2022;**50**:W541–50.

Perez-Arnaiz P, Dattani A, Smith V *et al.* *Haloferax volcanii*-a model archaeon for studying DNA replication and repair. *Open Biol* 2020;**10**:200293.

Pietilä MK, Roine E, Sencilo A *et al.* Pleolipoviridae, a newly proposed family comprising archaeal pleomorphic viruses with single-stranded or double-stranded DNA genomes. *Arch Virol* 2016;**161**:249–56.

Roine E, Kukkaro P, Paulin L *et al.* New, closely related haloarchaeal viral elements with different nucleic acid types. *J Virol* 2010;**84**:3682–9.

Roux S, Hallam SJ, Woyke T *et al.* Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. *eLife* 2015;**4**:e08490.

Sacher JC, Javed MA, Crippen CS *et al.* Reduced infection efficiency of phage NCTC 12673 on non-motile *Campylobacter jejuni* strains is related to oxidative stress. *Viruses* 2021;**13**:1955.

Schade SZ, Adler J, Ris H. How bacteriophage χ attacks motile bacteria. *J Virol* 1967;**1**:599–609.

Schiller H, Hong Y, Kouassi J *et al.* Identification of structural and regulatory cell-shape determinants in *Haloferax volcanii*. *Nat Commun* 2024;**15**:1414.

Schulze S, Adams Z, Cerletti M *et al.* The Archaeal Proteome Project advances knowledge about archaeal cell biology through comprehensive proteomics. *Nat Commun* 2020;**11**:3145.

Schuster H, Beyersmann D, Mikolajczyk M *et al.* Prophage induction by high temperature in thermosensitive DNA mutants lysogenic for bacteriophage lambda. *J Virol* 1973;**11**:879–85.

Schwarzer S, Rodriguez-Franco M, Oksanen HM *et al.* Growth phase dependent cell shape of haloarcula. *Microorganisms* 2021;**9**:231.

Sherman BT, Hao M, Qiu J *et al.* DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* 2022;**50**:W216–21.

Shevchenko A, Wilm M, Vorm O *et al.* A strategy for identifying gel-separated proteins in sequence databases by MS alone. *Biochem Soc Trans* 1996;**24**:893–6.

Shevchenko A, Wilm M, Vorm O *et al.* Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 1996;**68**:850–8.

Stern A. Sorek R The phage-host arms race: shaping the evolution of microbes *Bioessays* 2011;**33**:43–51.

Su Z-J, Wei Y-Y, Yin D *et al.* Effect of Sophora subprosrate polysaccharide on oxidative stress induced by PCV2 infection in RAW264.7 cells. *Int J Biol Macromol* 2013;**62**:457–64.

Tittes C, Schwarzer S, Pfeiffer F *et al.* Cellular and genomic properties of *Haloferax gibbonsii* LR2-5, the host of euryarchaeal virus HFTV1. *Front Microbiol* 2021a;**12**:625599.

Tittes C, Schwarzer S, Quax TEF. Viral hijack of filamentous surface structures in archaea and bacteria. *Viruses* 2021b;**13**:164.

Touchon M, Moura de Sousa JA, Rocha EP. Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Curr Opin Microbiol* 2017;**38**:66–73.

Turgeman-Grott I, Golan N, Neri U *et al.* A new archaeal virus that suppresses the transcription of host immunity genes. *bioRxiv* 2024, https://doi.org/10.1101/2024.02.12.579488.

Vizcaíno JA, Csordas A, del-Toro N *et al.* 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 2016;**44**:D447–56.

Wan Z, Goddard NL. Competition between conjugation and M13 phage infection in *Escherichia coli* in the absence of selection pressure: a kinetic study. *G3* 2012;**2**:1137–44.

Wang J, Liu Y, Liu Y *et al.* A novel family of tyrosine integrases encoded by the temperate pleolipovirus SNJ2. *Nucleic Acids Res* 2018;**46**:2521–36.

Wang X, Kim Y, Wood TK. Control and benefits of CP4-57 prophage excision in *Escherichia coli* biofilms. *ISME J* 2009;**3**:1164–79.

Williams KP. Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. *Nucleic Acids Res* 2002;**30**:866–75.

Wirth J, Young M. The intriguing world of archaeal viruses. *PLoS Pathog* 2020;**16**:e1008574.

Wolters M, Borst A, Pfeiffer F *et al.* Bioinformatic and genetic characterization of three genes localized adjacent to the major replication origin of *Haloferax volcanii*. *FEMS Microbiol Lett* 2019;**366**:fnz238.

Zegans ME, Wagner JC, Cady KC *et al.* Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J Bacteriol* 2009;**191**:210–9.

Zhang M, Zhang T, Yu M *et al.* The life cycle transitions of temperate phages: regulating factors and potential ecological implications. *Viruses* 2022;**14**:1904.

Received 29 November 2024; **revised** 20 February 2025; **accepted** 15 May 2025