environmental microbiology

Environmental Microbiology (2021) **00**(00), 00–00 doi:10.1111/1462-2920.15802

Virus diversity in metagenomes of a lichen symbiosis (*Umbilicaria phaea*): complete viral genomes, putative hosts and elevational distributions

Dominik Merges,[1,2*†] **Francesco Dal Grande**,[1*†] **Carola Greve**,[2] **Jürgen Otte**[1] **and Imke Schmitt**[1,2,3]

[1]*Senckenberg Biodiversity and Climate Research Centre, Frankfurt am Main, Germany.*

[2]*LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Frankfurt am Main, Germany.*

[3]*Department of Biological Sciences, Goethe Universität Frankfurt, Frankfurt am Main, Germany.*

Summary

Viruses can play critical roles in symbioses by initiating horizontal gene transfer, affecting host phenotypes, or expanding their host's ecological niche. However, knowledge of viral diversity and distribution in symbiotic organisms remains elusive. Here we use deep-sequenced metagenomic DNA (PacBio Sequel II; two individuals), paired with a population genomics approach (Pool-seq; 11 populations, 550 individuals) to understand viral distributions in the lichen *Umbilicaria phaea*. We assess (i) viral diversity in lichen thalli, (ii) putative viral hosts (fungi, algae, bacteria) and (iii) viral distributions along two replicated elevation gradients. We identified five novel viruses, showing 28%–40% amino acid identity to known viruses. They tentatively belong to the families *Caulimoviridae*, *Myoviridae*, *Podoviridae* and *Siphoviridae*. Our analysis suggests that the Caulimovirus is associated with green algal photobionts (*Trebouxia*) of the lichen, and the remaining viruses with bacterial hosts. We did not detect viral sequences in the mycobiont. Caulimovirus abundance decreased with increasing elevation, a pattern reflected by a specific algal lineage hosting this virus. Bacteriophages showed population-specific patterns. Our work provides the first comprehensive insights into viruses associated with a lichen holobiont and suggests an interplay of viral hosts and environment in structuring viral distributions.

Introduction

Viruses are obligate components of certain symbioses (Moran *et al*., 2005; Roossinck, 2015b; Hillman *et al*., 2018). They may impact the ecology of their hosts by affecting host phenotypic plasticity, initiating horizontal gene transfer, or expanding their host's environmental tolerances and ecological niche (Moya *et al*., 2000; Luis *et al*., 2007; Prasanna *et al*., 2010). For example, the presence of a laterally transferred viral gene in aphids enables a rapid development of wings in response to high population densities (Parker and Brisson, 2019). In a three-way symbiosis between a plant, a fungus and a virus, the virus confers thermal tolerance to the system and facilitates survival of the plant in soils with temperatures >50°C (Luis *et al*., 2007).

The geographic distribution of viruses in natural populations is only beginning to be investigated. Recent advances in DNA sequencing have enabled mining for viral sequences from complex metagenomic samples collected over the entire distribution of a viral host species of interest (Massart *et al*., 2014; Hayes *et al*., 2017; Simmonds *et al*., 2017). Broad scale, multi-population sampling combined with metagenomic sequencing now permits biogeographic studies of viruses and their hosts (Shi *et al*., 2017; Bergner *et al*., 2020). So far, however, investigations of this type have been restricted to marine environments (Suttle, 2005; Brum and Sullivan, 2015; Chow and Suttle, 2015; Coutinho *et al*., 2017; Breitbart *et al*., 2018), and animal hosts (Remnant *et al*., 2017; Shi *et al*., 2017; Greninger, 2018).

The integration of an ecological framework, such as diversity analyses across gradients, into viral metagenomic studies can address questions which have been mostly neglected in virology but are key to advance our understanding of viruses and how they interact with their environment (Sommers *et al*., 2021). A replicated elevational gradient setup is highly relevant to study a species' response to environmental change, in particular

Received 22 January, 2021; accepted 1 October, 2021. *For correspondence. E-mail dominik.merges@senckenberg.de; francesco.dalgrande@senckenberg.de; Tel. +49 69 7542-1856; Fax. +49 69 7542-1800. †These two authors contributed equally to this work and are considered joint first authors.

to climate. The assessment of taxonomic diversity is the first step to then move onto functional diversity. Our study covers two of the three basic axes of biodiversity: taxonomic and structural (community structure). Taken together these represent an essential basis to then understand functional trait diversity (e.g. Hillebrand *et al*., 2018). Ultimately the integration of an ecological framework in viral metagenomics could broadly expand our knowledge on ecosystem-level effects of viruses (Roux *et al*., 2021; Sommers *et al*., 2021).

It is presently unknown whether viruses contribute to the remarkable environmental tolerances of lichen-forming fungi. Lichens occur on all continents and in all climate zones, including Antarctica (Kappen, 2000). Some species have distributions across broad ecological conditions and show a wide range of environmental tolerances (Kappen, 2000; Werth and Sork, 2014; Singh *et al*., 2017). Desiccation tolerance and UV tolerance are typical traits of the lichen consortium (Grube and Berg, 2009). Lichens are complex, symbiotic organisms consisting of a primary fungal species (mycobiont), one or more photosynthetic partners (photobionts), and additional, highly diverse, prokaryotic and eukaryotic communities (Arnold *et al*., 2009; Grube *et al*., 2009; Bates *et al*., 2012; Fernández-Mendoza *et al*., 2017). The diversity of photobionts is often significantly correlated with geography and ecology, and may thus be involved in environmental specialization (Peksa and Škaloud, 2011; Vargas Castillo and Beck, 2012; Dal Grande *et al*., 2018). Furthermore, bacterial communities associated with lichens respond to photobiont type and environment (Hodkinson *et al*., 2012), and the mycobiont's genome can be differentiated within species along elevation (Dal Grande *et al*., 2017). Hence, lichens can be thought of as dynamic consortia that respond to environmental drivers with changes in genome composition and community re-arrangements within the holobiont. The potential role of viruses, as mobile genetic elements, in shaping genomes and microorganismal communities associated with lichens remains hidden.

Little is known about viral diversity associated with lichens. So far, only two viruses have been reported in lichen mycobionts, i.e. *Chrysothrix chrysovirus* 1 and *Lepraria chrysovirus* 1 (Petrzik *et al*., 2019). Similarly, in lichen photobionts, only a remnant of a large double-stranded DNA virus insertion was reported in *Asterochloris*, and phylogenetically placed in the Phycodnaviridae, a family of viruses that infect green algae (Armaleo *et al*., 2019). Additionally, two viruses have been reported in the common photobiont genus *Trebouxia*: a negative-strand RNA virus phylogenetically close to plant cytorhabdoviruses, and apple mosaic virus (Ilarvirus, Bromoviridae) (Petrzik *et al*., 2014), and the cauliflower mosaic virus (CaMV, Caulimoviridae), a circular, double-stranded DNA virus known from herbaceous plants (Petrzik *et al*., 2015). Surprisingly, very few viruses associated with bacteria on lichens have been reported so far, with only one study reporting the presence of bacteriophages in the genus *Peltigera* (Garg *et al*., 2016). It is currently unknown whether viruses (i) are inherent components across populations of lichen mycobionts, photobionts, or bacteria associated with lichens, and (ii) whether viruses contribute to the lichen holobiont's environmental tolerance.

Here we used a metagenomic approach to assess the diversity and distribution of lichen-associated viruses across 11 *Umbilicaria phaea* populations along two elevational gradients in North America. Samples ($n = 550$) were collected from the Sierra Nevada and Mount San Jacinto (California), covering almost 2000 elevational meters and pronounced climatic differences. We address the following research questions: (i) which viruses can be found in the *U*. *phaea* lichen symbiosis? (ii) which components of the lichen (i.e. mycobiont, photobiont) or lichen associated bacteria harbour viral sequences? and (iii) are there environmental and host-specific patterns of viral distributions in populations of *U*. *phaea* along elevational gradients?

Material and methods

Study site and sample collection

We sampled 11 *U*. *phaea* populations along two elevational gradients in California, USA (Table S1). The gradients are spatially separated by approximately 700 km. We collected fragments of 50 individuals each, at four populations along the Sierra Nevada gradient (38.084, −120.484), and at seven populations along the Mt. San Jacinto gradient (33.435, −116.484; Table S1). We collected the thalli using sterile tweezers and placed them individually into sterile paper bags until further processing. Thalli were cleaned from bark, moss, insects and other visible contaminations under a dissecting microscope prior to DNA isolation (Hodkinson and Lutzoni, 2009; Hodkinson *et al*., 2012; Aschenbrenner *et al*., 2014; Ramírez-Fernández *et al*., 2014; Grube *et al*., 2015; Cernava *et al*., 2017; Graham *et al*., 2018). We additionally collected two whole lichen thalli, one low altitude individual from the Sierra Nevada population *Uph16* and one from the high altitude population *Uph19*, for the reconstruction of reference genomes using PacBio Sequel II data (for detailed description please see Supporting Information 1 and Table S1).

DNA extraction for population pooled sequencing

Genomic DNA was extracted separately from each fragment from all populations using a CTAB-based method

(Cubero and Crespo, 2002; Dal Grande *et al*., 2017). In brief, 5 ml of extraction buffer (1% wt./vol. CTAB; 1 M NaC1; 100 mM Tris; 20 mM EDTA; 1% wt./vol. polyvinyl polypyrolidone) was added to the ground material and mixed by inverting them several times and then heated for 30 min at 70°C before adding one volume of chloroform:isoamyl alcohol (24:1 vol./vol.) (Cubero and Crespo, 2002). The resulting product was mixed by inverting the tube and centrifuged for 5 min 10 000*g* (Cubero and Crespo, 2002). Afterwards, the upper aqueous phase was transferred to a new tube and the lower layers were discarded (Cubero and Crespo, 2002). Two volumes of precipitation buffer (1% wt./vol. CTAB; 50 mM Tris-HCl; 10 mM EDTA; 40 mM NaC1) were added to the transferred aqueous phase and mixed for 2 min by inversion (Cubero and Crespo, 2002). Subsequently, the mixture was centrifuged for 15 min at 13 000*g* and the pellet was resuspended in 350 µl of 1.2 M NaCl, to which one volume of chloroform:isoamylalcohol (24:1) was added (Cubero and Crespo, 2002). This was mixed and centrifuged for 5 minutes at 10000g at room temperature (Cubero and Crespo, 2002). The upper phase was transferred to a new tube and 0.6 volumes of isopropanol was added (Cubero and Crespo, 2002). The product was mixed by inversion and incubated at −20°C for 15 min (Cubero and Crespo, 2002). Afterwards, the final pellet was produced by centrifugation for 20 min at 13 000 g at 4°C (Cubero and Crespo, 2002). The final pellet was washed with 1 ml of 70% ethanol and collected by centrifugation for 3 min at 13 000 g at 4°C (Cubero and Crespo, 2002). The pellet was dried at 50°C prior to resuspension in TE buffer (10 mM Tris pH 7.4, 1 mM EDTA) (Cubero and Crespo, 2002).

Furthermore, we created a pooled sample for each population containing equal amounts of DNA from each sample [i.e. Pool-seq; (Dal Grande *et al*., 2017)]. Novogene (Cambridge, United Kingdom) performed the library preparation (200–300 bp insert size), sequencing on an Illumina HiSeq2000 with 150 bp paired-end chemistry at ∼90× coverage per population (for the mycobiont fraction), and delivered sequences trimmed of tags and adaptors.

DNA extraction for genomic sequencing

Genomic DNA for genome sequencing was extracted from dry thallus material of *Uph*16 and *Uph*19 samples according to Mayjonade *et al*. (2016) (for detailed description please see Supporting Information 1).

PacBio library preparation and sequencing

Two SMRTbell libraries were constructed according to the manufacturer's instructions of the SMRTbell Express Prep kit v2.0 following the Low DNA Input Protocol (Pacific Biosciences, Menlo Park, CA). Total input DNA for *Uph16* and *Uph19* samples was approximately 140 and 600 ng respectively. Ligation with T-overhang SMRTbell adapters was performed at 20°C overnight. Following ligation, the SMRTbell library was purified with an AMPure PB bead clean up step with 0.45× volume of AMPure PB beads. Subsequently, a size-selection step with AMPure PB Beads was performed to remove short SMRTbell templates <3 kb. For this purpose, the AMPure PB beads stock solution was diluted with elution buffer (40% vol./vol.) and then added to the DNA sample with 2.2× volume. The size and concentration of the final libraries were assessed using the TapeStation (Agilent Technologies) and the Qubit Fluorometer with Qubit dsDNA HS reagents Assay kit (Thermo Fisher Scientific, Waltham, MA). Sequencing primer v4 and Sequel® II Polymerase 2.0 were annealed and bound respectively, to each SMRTbell library. SMRT sequencing was performed on the Sequel System II with Sequel II Sequencing Kit 2.0 in 'continuous long read' (i.e. CLR) mode, 30 h movie time with NO preextension and Software SMRTLINK 8.0. One SMRT Cell was run for each sample.

De novo *assembly of PacBio metagenomic sequence reads*

We used the continuous long reads (i.e. CLR reads) from the Pacbio Sequel II CLR run (i) to process them into highly accurate consensus sequences (i.e. HiFi reads) for genome assemblies of the most abundant lichen components (i.e. the myco- and photobiont) and (ii) for direct assembly of remaining raw CLR reads to explore the microbial fraction of the lichen holobiont. HiFi reads were retrieved from raw CLR reads with the PacBio tool CCS v5.0.0 using default parameters (https://ccs.how). Metagenomic sequence reads were assembled into contigs using the assembler metaFlye v2.7 (Kolmogorov *et al*., 2019). The assembled contigs were scaffolded with SSPACE-LongRead v1.1 (Boetzer and Pirovano, 2014). The received scaffolds were taxonomically binned via blastx using DIAMOND (--more-sensitive --frameshift 15 –range-culling) on a custom database and the MEGAN6 Community Edition pipeline (Buchfink *et al*., 2014; Huson *et al*., 2007). The completeness of the reconstructed fungal and algal genomes was estimated using Benchmarking Universal Single-Copy Orthologues (BUSCO) analysis in BUSCO v4 (Simão *et al*., 2015). The resulting HiFi read assemblies were used to remove the fungal and algal proportion from the CLR reads, to allow the assembly of the bacterial fraction of the lichen holobiont. These filtered CLR reads were assembled and scaffolded as described above, with an additional polishing step in between, using the PacBio tool GCpp v2.0.0 with default parameters (github.com/PacificBiosciences/gcpp).

Assessing virus presence in the assembled metagenomic sequence reads

Presence of known lichen viruses. All available genomes of viruses putatively associated with lichens ($n = 5$) or with closely related genera were identified from the literature ($n = 514$) (Petrzik *et al*., 2014, 2015, 2019; Armaleo *et al*., 2019; Greshake Tzovaras *et al*., 2020). The genomes of these 519 viral candidates were retrieved from GenBank at the National Center for Biotechnology Information (ncbi.nlm.nih.gov/geo/). The mycobiont and photobiont fractions of the assembled HiFi metagenomic reads were searched for insertion of viral genomes using the nucleotide search of the Basic Local Alignment Search Tool v2.10.0 (BLASTn) with an *e*-value 1*e*-6 (Altschul *et al*., 1990).

Identification of novel lichen viruses. The assembled CLR metagenomic sequence reads were screened for the presence of viral scaffolds with VIBRANT v1.2.0 (Kieft *et al*., 2020). We used VIBRANT based on its strength in recovering full viral genomes and identifying integrated provirus sites. Recovering integrated proviruses then allowed us to retrieve information on the site coordinates within scaffolds, and enabled us to analyse phage and host sequence read distribution. VIBRANT advances previous virome detection tools (e.g. VirSorter, VirFinder, MARVEL, VirusSeeker) by leveraging neural networks of protein annotation signatures from non-reference-based similarity searches with HMMs from three databases: Kyoto Encyclopaedia of Genes and Genomes KoFam, Pfam (v32) and Virus Orthologous Groups (release 94, vogdb.org), and implementing curation steps to validate predictions based on scoring of detected viral sequences (Kieft *et al*., 2020). Thereby VIBRANT outperformed VirFinder, VirSorter and MARVEL in the ability to maximize virus recovery and minimize false discovery (Kieft *et al*., 2020). Subsequently, the completeness of viral scaffolds was estimated with CheckV v0.6.0 (Nayfach *et al*., 2020). Taxonomic classification of scaffolds was done with the NCBI RefSeq database in Kaiju v1.7.3 (Menzel *et al*., 2016).

Identification of putative viral hosts.

i. Literature research: References associated with the viral taxonomic assignments from NCBI were screened for published host associations.
ii. Taxonomic classification of host sequence: For viral fragments hosted within larger scaffolds (e.g. prophages) the flanking regions (i.e. DNA of host sequences) were extracted and taxonomically assigned with an NCBI nucleotide BLAST (Altschul *et al*., 1990). For the extraction of the flanking regions from scaffolds containing viral sequences (i.e. DNA of host sequences), the integrated prophage coordinates were retrieved from VIBRANT output 'integrated_prophage_coordinates_scaffolds.tsv'. The coordinates were used to hardmask (i.e. convert bases to N) the integrated viral sequences within the scaffolds using bedtools v2.28.0 (Quinlan and Hall, 2010).
iii. Blasting of viral scaffolds against the HiFi metagenomic sequencing read assemblies:

The identified viral scaffolds were blasted against the mycobiont and photobiont scaffolds assembled from the HiFi reads for the *Uph16* and *Uph19* samples respectively.

Viral distribution across U. phaea *populations.* We filtered the pool-seq data for reads shorter than 80 bp, reads with N's and reads with average base quality scores less than 26 along with their pairs, and discarded them. We mapped the trimmed paired-end reads of each pool to the database of the identified viral scaffolds using bowtie2 v2.4.1 (Langmead and Salzberg, 2012), using the flags: --very-sensitive-local, --no-mixed, --no-unal, --no-discordant. The number of mapped reads was counted per sample and normalized by dividing the number of mapped reads by the total read number of the respective sample to account for differences in sequencing depth. Normalized read numbers and linear models (normalized reads $\sim$ elevation) were fitted and plotted in R v3.6.1 (R Core Team, 2019).

Host abundance across U. phaea *populations*

Inferring photobiont abundance. Each pool-seq dataset was mapped to the dominant photobiont ITS haplotypes present at the gradients (Rolshausen *et al*., 2020) using bowtie2 (flags: --very-sensitive-local, --no-mixed, --no-unal, --no-discordant). For each population, the mapped reads were extracted and assembled using metaSPAdes v3.14.1 (Nurk *et al*., 2017). The assembled contigs were then screened for photobiont full ITS sequences using ITSx (Bengtsson-Palme *et al*., 2013). The retrieved ITS sequences were blasted against the dominant *Trebouxia* haplotypes and identified based on 100% identity (Altschul *et al*., 1997; Rolshausen *et al*., 2020). The abundance of the recovered haplotypes across the populations was then assessed by mapping and plotting as described above.

Inferring bacterial abundance. The extracted flanking regions (i.e. DNA of host sequences) were mapped using bowtie2 with stringent settings (flags: --very-sensitive-local, --no-mixed, --no-unal, --no-discordant) against the pool-seq data and plotted as described above. Similarity in abundance patterns of viruses and their putative hosts

were assessed with correlation tests using Pearson's r in R v3.6.1 (R Core Team, 2019).

Results

HiFi metagenomic sequencing reads of mycobiont, photobiont and microbial communities

We reconstructed metagenomic sequences from a low altitude (*Uph16*, 631 m a.s.l.) and a high altitude (*Uph19*, 2036 m a.s.l.) specimen of *U. phaea* (Tables 1 and 2). Sequence output and quality are summarized in Table 1. Both mycobiont genomes were of the expected size, highly contiguous, complete and sequenced to a mean read coverage of 1042× and 5577× respectively (Table 2). The *Trebouxia* photobiont genomes were incomplete and fragmented and did not match the expected size of around 70 Mb [Table 2; (Greshake Tzovaras *et al*., 2020)].

The bacterial metagenome fraction had a total length of 35.2 Mb comprising 550 scaffolds with an N50 of 0.14 Mb (Table 2). Furthermore, a viral genome of a total length of 0.11 Mb containing two scaffolds with an N50 of 0.1 Mb was recovered (Table 2). Of these two scaffolds, scaffold 209 (S209) contained a complete viral genome with a total genome length of 101 900 bp, 100% CheckV completeness and a mean read coverage of 167× (Tables 3 and 4, *U. phaea* siphovirus-like 1). The mycobiont and photobiont fractions assembled from the metagenomic sequencing reads did not contain any insertion of previously known viruses. Scaffold 105, from the photobiont fraction of the low altitude *Uph16* sample, partly mapped to the newly constructed viral scaffold S3586 recovered (for details: see next paragraph).

Viral genomes retrieved from assembled CLR metagenomic sequence reads

We screened the assembled metagenomic sequences from the CLR datasets of the low altitude *Uph16* and high altitude *Uph19* samples for viral genomes. From these data, we identified four additional viruses, and tentatively placed them in the *Caulimoviridae*, *Myoviridae*, *Podoviridae* and *Siphoviridae*. *Caulimoviridae* are associated with vascular plants, all other viral families with bacteria. All viruses identified were novel, exhibiting only 28%–40% amino acid identity to previously identified viruses (Table 3). Genome statistics of the viruses found in the low altitude *Uph16* and high altitude *Uph19* samples are given in Table 3. *Siphoviridae* scaffold 1070 (termed *U. phaea* siphovirus-like 2, Table 4) was classified as lysogenic bacteriophage, containing 19 predicted genes encoding phage specific proteins, such as phage tail tube protein (PF10618.9), Phage terminase (PF03354.15) and phage portal protein (PF04860.12), for detailed information see Supplemental Table S3. Caulimovirus scaffold 3586 (termed *U. phaea* caulimovirus-like 1, Table 4) was predicted to contain 13 genes, with two genes encoding putative proteins that matched P03556 (POL_CAMVD) the ORF V *of CaMV*, *a virus that infects plants*, *and contains a strong promoter (35 s)*, *which is used in biotechnology [e.g*. (Amack and Antunes, 2020]. Myovirus scaffold 18 (termed *U. phaea* myovirus-like 1, Table 4) and Podovirus scaffold 27 (termed *U. phaea* podovirus-like 1, Table 4) had predicted host boundaries at 5′ and 3′ ends and thereby were classified as proviruses. The host sequences were taxonomically placed within clades of Gram-negative bacteria. *Umbilicaria phaea* myovirus-like 1 (S18) was hosted within a bacterial sequence most closely matching the genus *Roseomonas*, a clade of widely distributed bacteria. *Umbilicaria phaea* podovirus-like 1 was found within the bacterial sequence taxonomically placed in the *Sphingomonas* group. For coverage information of the putative hosts across the pooled sequencing data please see Supplementary Table 4.

Viral abundance distributions along gradients

The normalized read number of *U. phaea* caulimovirus-like 1 scaffold 3586 showed a sharp decline with increasing elevation (Fig. 1A), whereas the four segments of bacteriophage origin showed localized abundance trends (Fig. 1B and C, Fig. S1). The Caulimovirus normalized read numbers showed a significant decrease with increasing elevation ($p < 0.05$; Fig. 1A). The *U. phaea* myovirus-like 1 (S18) showed a gradient-specific response to elevation, with a non-significant trend of increasing abundance with increasing elevation along the Sierra Nevada gradient (Fig. 1B, blue circles). Similarly, the normalized read number of *U. phaea* podovirus-like 1 S27 showed a non-significant trend towards a gradient-

Table 1. Metrics of the PacBio Sequel II HiFi read yield.

Sample	Sequence output (Gb)	n CLR reads	n HiFi reads	Yield in %	Total size (Gbp)	Mean read length (bp)	Median read length (bp)
Low altitude *U. phaea Uph16*	108	3 750 852	214 402	5.72	1.46	6792	6060
High altitude *U. phaea Uhp19*	564	7 240 242	1 619 047	22.36	14.04	8674	7294

Table 2. Metrics of the assembled HiFi metagenomic sequence reads.

Sample	Assembly method	Taxonomic classification	Number of scaffolds	Total length (Mb)	N50 (Mb)	BUSCO completeness	Mean coverage
Low altitude	metaFlye	Fungal	47	35.6	1.26	98.5	1042×
U. phaea Uph16		Algal	37	1.09	0.028	/	481×
		Bacterial	14	0.48	0.04	/	159×
High altitude	metaFlye	Fungal	38	35.1	1.57	98.5	5577×
U. phaea Uph19		Algal	884	97.2	0.57	94.9	248×
		Bacterial	550	35.2	0.14	/	482×

specific response along the Sierra Nevada gradient (Fig. 1C, blue circles). The abundances of *U. phaea* siphovirus-like 1 (S209, Fig. S1a) and 2 (S1070, Fig. S1b) showed a trend towards a localized population-specific response, with high read numbers only within two (S209 in *Uph18* and *Uph19*) and one population (S1070 in *Uph16*) respectively (Fig. S1a and b; Table S1).

Elevational distribution of viral hosts

Three full ITS sequences of the lichen's photobionts could be reconstructed from the pool-seq dataset (Fig. S2; Table S1). One ITS sequence corresponded to the *Trebouxia* haplotype *Hap01* dominant at low altitudes, and one ITS sequence to the *Trebouxia* haplotype *Hap03* dominant at high altitudes, as identified in a previous study (Rolshausen *et al*., 2020). The blast search of the third ITS sequence against the *Trebouxia* haplotypes did not reveal a correspondence to any known *Trebouxia* haplotype. The distribution of the *U. phaea* caulimovirus-like 1 (S3586, Fig. 1A) was significantly correlated with the distribution of the dominant low altitude haplotype *Hap01* (Pearson's r 0.97, $p < 0.001$; Fig. 1D). The distribution of normalized read numbers of the *U. phaea* myovirus-like 1 (S18) was correlated with the distribution of normalized read numbers of its flanking host regions (Pearson's r 0.98, $p < 0.001$; Fig. 1B and E). The distribution of *U. phaea* podovirus-like 1 (S27), for which the flanking host regions could be recovered, showed a less correlated abundance pattern with its host across the elevational gradient (Pearson's r 0.72, $p = 0.01$; Fig. 1C and F).

Discussion

Although viruses are ubiquitous in nature and can play vital roles in symbiosis, viral diversity and distribution in populations of a single host species are still largely unknown. In this study, we take an in-depth look at viral diversity of the holobiont of the lichen *U. phaea* and trace the discovered viruses in lichen populations along two replicated elevation gradients. The discovered viruses are not shared among the two components of the lichen symbiosis (mycobiont, photobionts) and their associated bacteria, suggesting they are rather host specific. This finding also suggests the lack of virus-mediated inter-kingdom horizontal gene transfer in the studied lichen. We report, for the first time, viruses associated with bacteria from lichen thalli based on metagenomic data. In the following, we discuss our findings with regard to the different partitions of the *U. phaea* lichen symbiosis: mycobiont, photobiont and bacteria associated with the lichen thalli.

Fungal viruses

Mycoviruses are commonly found in fungi (Göker *et al*., 2011; Ghabrial *et al*., 2015; Hillman *et al*., 2018), and can be asymptomatic, or causing phenotypic changes (Nuss, 2005; Simmonds *et al*., 2017). Viruses are present in the *Eurotiomycetes* (Kotta-Loizou and Coutts, 2017) a class closely related to the mostly lichenized *Lecanoromycetes*. Recently, two viruses have been isolated from the lichenized fungi *Chrysothrix chlorina* and *Lepraria incana* (Petrzik *et al*., 2019). Thus, we were surprised not to detect any viruses in the deeply sequenced and highly contiguous fungal genomes of *U. phaea*. A methodological reason for this could be lack of detection due to incomplete viral sequence archives. Viral detection relies on sequence similarity searches such as BLAST and HMMER (as implemented in VIBRANT & CheckV) against currently known viral sequences. Therefore, viruses that are too divergent from sequences available in databases may remain unnoticed, and the apparent absence of viral sequences in *U. phaea* does not exclude the possibility, that viruses are actually present. This is in line with the assumption that the largest fraction of viral diversity remains undiscovered to date (Roossinck, 2015a; Marzano and Domier, 2016; Roux *et al*., 2019). Furthermore, the current study is based on DNA and thus cannot inform about the diversity of mycobiont viruses potentially present in RNA.

Table 3. Quality metrics of viral genomes and their putative hosts.

Sample	Scaffold	Genome length (bps)	Viral genes	aai id	aai af	Completeness	Genome structure	Mean coverage	NCBI taxon identifier of best match	Virus identity (best match)	Order	Family	Putative host	Host sequence	Host identity (best match)
Low altitude *U. phaea Uph16*	S1070	41 475	17	36.4	16.4	97.18	Linear	18×	1 147 150	Enterobacteria phage mEp235	Caudovirales	Siphoviridae	Bacteria	na	na
	S3586	5901	2	28.7	33.8	100	Linear	730×	930 168	Sweet potato collusive virus	Ortervirales	Caulimoviridae	Plants	na	na
High altitude *U. phaea Uph19*	S18	46 773	22	39.5	24.3	100	Linear	346×	2 733 083	Tigrvirus	Caudovirales	Myoviridae	Bacteria	CP015583.1	*Roseomonas gilardii* strain U14-5
	S27	36 810	25	35	15.1	100	Linear	608×	179 237	*Sinorhizobium* phage PBC5	Caudovirales	Podoviridae	Bacteria	CP042306.1	*Sphingomonas* sp. HKS19
	S209	101 900	33	46.5	20.6	100	Linear	167×	2 733 152	Bertelyvirus	Caudovirales	Siphoviridae	Bacteria	na	na

Average amino acid identity over the contigs coding length was abbreviated as 'aai id', the alignment fraction of the contig was abbreviated as 'aai af'.

Table 4. Proposed names and GenBank accession numbers for viral nucleotide sequences.

Name	Scaffold	GenBank Accession number
Umbilicaria phaea podovirus-like 1	S27	OK513178
Umbilicaria phaea siphovirus-like 1	S209	OK513179
Umbilicaria phaea siphovirus-like 2	S1070	OK513180
Umbilicaria phaea myovirus-like 1	S18	OK513181
Umbilicaria phaea caulimovirus-like 1	S3586	OK513182

Algal viruses

Algal viruses are highly diverse and associated with a large number of taxonomically unrelated hosts (Short, 2012). A few algal viruses have been reported from lichen photobionts, such as *Asterochloris glomerata*, *Chlorella* spp. and *Trebouxia* sp. (Petrzik *et al*., 2014, 2015; Armaleo *et al*., 2019). Our study is the first to show geographic distributions of an algal virus in a terrestrial ecosystem. The detected *U. phaea* caulimovirus-like 1 significantly decreases with increasing elevation along both studied gradients. This pattern is mirrored by the distribution of one of the *Trebouxia* lineages (*Hap01*) found in the *U. phaea* populations. One explanation for this elevational pattern could be that environmental factors structure the occurrence of the alga (Dal Grande *et al*., 2018; Rolshausen *et al*., 2018, 2020), and the virus tracks the algal host as a passenger. Petrzik *et al*. (2015) found that the CaMV (family Caulimoviridae) present in *Trebouxia aggregata* (isolated from the lichen *Xanthoria parietina*) was very faithful to its host, and stayed with it under cultivation conditions for more than 70 years without causing any detrimental damage to the algal culture. If Caulimoviruses are indeed rather benign to trebouxioid hosts, then one could speculate that the Caulimovirus of the present study is tracking host abundance in form of a chronic infection, rather than controlling host abundance, e.g. by displaying differential pathogenicity along the climatic gradient. Our results, together with the report by Petrzik *et al*. (2015) of CaMV in *T. aggregata*, photobiont of the globally distributed lichen species *X. parietina*, suggest that lichen photobionts are common, natural hosts for these plant-associated viruses.

The *U. phaea* caulimovirus-like 1 identified in our study was predicted to encode for two genes that matched to the ORF V *of CaMV*, *a common* pathogen that infects plants in the *Brassicaceae* family and is transmitted by aphids as animal vectors (Amack and Antunes, 2020; Bak and Emerson, 2020). The *CaMV* encodes a transcriptional promoter (*CaMV* 35s promoter) that is used extensively to overexpress genes in plants, based on its insensitivity to endogenous plant signals when used as a

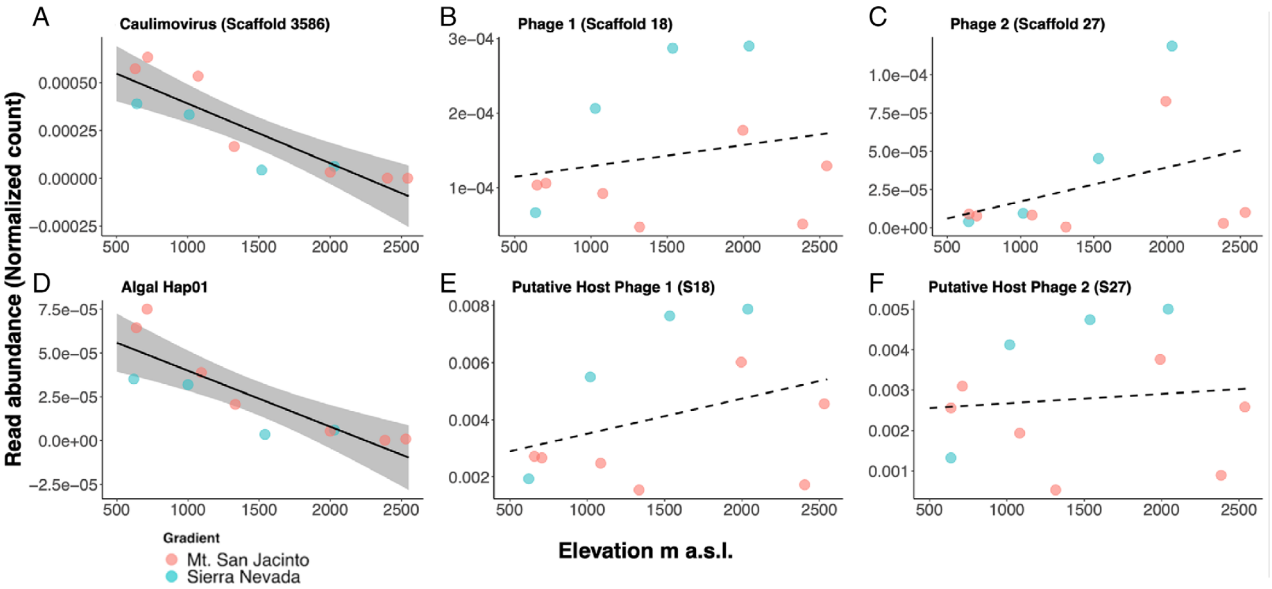

Fig. 1. The abundance of viruses and their putative hosts across the elevational gradients. *Umbilicaria phaea* caulimovirus-like 1 (A) and algal Haplotype Hap01 (D) abundance patterns were highly correlated (Pearson's r 0.97, $p < 0.001$) and decreased significantly with increasing elevation. Phage 1 (*U. phaea* myovirus-like 1, Scaffold 18) (B) and its putative host (E) abundance patterns were highly correlated (Pearson's r 0.98, $p < 0.001$) and showed a gradient specific response to elevation. Phage 2 (*U. phaea* podovirus-like 1, Scaffold 27) (C) and its putative host (F) distributions were less strongly correlated (Pearson's r 0.72, $p = 0.01$) and showed a gradient specific response to elevation. Linear models of read abundances were fitted. Raw data are shown in red (Mt. San Jacinto) and blue (Sierra Nevada) circles. Dashed lines show non-significant effect of elevation on viral abundance ($p > 0.05$). [Color figure can be viewed at wileyonlinelibrary.com]

heterologous promoter (Wu *et al*., 2014; Amack and Antunes, 2020). As a consequence, the CaMV 35S promoter is present in approx. 60% of all genetically modified crops currently grown worldwide, thereby being among the most widely employed genetic tools in agricultural biotechnology (Amack and Antunes, 2020). Given the presence of two genes that matched the ORF V of CaMV in our study, this genome could be explored for alternative heterologous promoters.

Most algal viruses known to date have narrow host ranges and produce species-specific or even strain-specific infections (Short, 2012). Our data also suggest a certain degree of host specificity, with the detected *U. phaea* caulimovirus-like 1 apparently not spreading to other algal lineages associated with *U. phaea*. Indeed, we can corroborate the significant correlation between the *U. phaea* caulimovirus-like 1 and *Trebouxia Hap01* with transcriptomic data from algal haplotype cultures, where we find the *U. phaea* caulimovirus-like 1 to be abundant in *Hap01* cultures (unpublished data; Supplementary Table S2; Fig. S1). This is interesting because single lichen thalli are long-lived structures that commonly harbour several different lineages of trebouxioid photobionts (Casano *et al*., 2011; Dal Grande *et al*., 2014; Paul *et al*., 2018). The close physical contact of algal strains over extended periods of time should theoretically allow high encounter rates and provide ample opportunity for viral transmission, yet the present study supports the view that cross strain infections are not common.

Bacterial viruses

Bacteriophages are known to play critical roles in bacterial mortality and horizontal gene transfer (Suttle, 2005; Manrique *et al*., 2016; Petrzik *et al*., 2019). Bacterial diversity associated with lichens is high, and some components of lichen-associated bacterial communities have been suggested to contribute functions to the symbiosis, such as stress amelioration, pathogen defence and detoxification (Grube *et al*., 2009, 2015; Cernava *et al*., 2017). The presence of viruses in the lichen-associated bacteria has not been assessed. Here, we report four novel bacterial viruses, phylogenetically placed in the families *Myoviridae*, *Podoviridae* and *Siphoviridae* (*Caudovirales*). The three families contain tailed bacterial viruses parasitizing bacteria (Breitbart *et al*., 2004; Lavigne *et al*., 2011). These viral families show a broad phylogenetic distribution, being present in 140 prokaryote host genera spread across most branches of the bacterial tree (Lavigne *et al*., 2011). Their host specificity ranges from host generalists to host specialists, with some viruses infecting hosts of multiple genera, but most are specific for particular groups of closely related host species (McGrath *et al*., 2004; Lavigne *et al*., 2011; Paez-Espino *et al*., 2016). The Caudovirales

belong to the main groups of dsDNA phages which are detected in environmental samples across most ecosystems (Wommack *et al*., 1999; Breitbart *et al*., 2004; Lavigne *et al*., 2011). Our result contributes to the understanding of the phylogenetic breadth of bacteriophage hosts and supports the lichen environment as a suitable habitat for members of the Caudovirales.

The four detected bacteriophages show different distribution patterns along the gradients. *Umbilicaria phaea* myovirus-like 1 (S18, Phage 1) shows no elevational structure, but correspondence to the distribution of its bacterial host. The bacterial host was placed in the genus *Roseomonas*, a widely distributed bacterial clade. *Roseomonas* spp. occur across vastly different environments from plants, water, arctic tundra soils and human blood (Gallego *et al*., 2006; Gundlapally & Garcia-Pichel, 2006; Jiang *et al*., 2006; Y. Q. Zhang *et al*., 2008). Aside from the pathogenicity reported of some *Roseomonas* strains for humans, little is known about functions in natural systems or about viruses infecting them. *Umbilicaria phaea* podovirus-like 1 (S27, Phage 2) shows elevational structure in only one of the gradients (Sierra Nevada) and has no correspondence to the distribution of its bacterial host. The bacterial host is within the *Sphingomonas* group, which contains bacteria often associated with plants, where they can play vital roles in plant abiotic stress tolerance (Kim *et al*., 1998; Khan *et al*., 2014, 2017; Luo *et al*., 2019; Asaf *et al*., 2020). Phages are commonly found in *Sphingomonas* and exhibit a narrow to broad host range (Nakayama *et al*., 2009). The role of phages in the *Sphingomonas* mediated stress tolerance for plants is unknown. Members of the genera *Roseomonas* and *Sphingomonas* have been detected a wide range of lichen taxa (Hodkinson *et al*., 2012; Jiang *et al*., 2017; Noh *et al*., 2020, 2021). The two Siphoviruses reported in our study *U. phaea* siphovirus-like 1 (S209, Phage 3) and 2 (S1070, Phage 4) showed more localized patterns. *Umbilicaria phaea* siphovirus-like 1 (Phage 3) has occurrences only in the two highest sites of the Sierra Nevada gradient and is otherwise rare. *Umbilicaria phaea* siphovirus-like 2 (S1070, Phage 4) shows a highly localized occurrence in only a single population. These patterns suggest random occurrences of these viruses at the specific site and time of sampling.

Overall, the bacteriophages detected in this study showed stochastic distributions. We cannot distinguish whether the bacterial hosts of these viruses belong to the lichen microbiome in the strict sense or whether they are environmental bacteria. A future study, based on individual-based resequencing, assessing the abundances of bacteriophages would help clarifying (i) whether the observed pattern in the Pool-seq data is related to low coverage of the microbiome of individual lichens and (ii) what is the actual frequency of the phages within the populations.

Drivers of viral elevational distribution

One of the viruses (*U. phaea* caulimovirus-like 1) we found showed abundance patterns that strongly correlated with elevation. This pattern could be explained by a change in host diversity, alternative host species, or environmental conditions co-varying with elevation. For putative viral hosts in lichen-associated communities (i.e. mycobiont and photobiont), spatial heterogeneity has been found to be high, showing pronounced turnover in community composition within the same gradient, between different gradients, or over large latitudinal scales (Singh *et al*., 2017; Dal Grande *et al*., 2018; Rolshausen *et al*., 2018, 2020). In bats, a decline in viral richness along an elevational gradient was mainly linked to a decline of hosts towards higher elevations (Bergner *et al*., 2020). Accordingly, shifts in viral diversity and distribution within lichens could be reflecting the turnover of their hosts along the elevational gradient. For example, in our study, the algal-associated *U. phaea* caulimovirus-like 1 co-occurred almost exclusively with the photobiont *Hap01* found at low elevations. However, an interplay between host turnover and direct effects of environmental factors on viruses could be possible, as environmental factors may directly affect viral persistence (Guernier *et al*., 2004). For example in Antarctic soils, declining viral richness with increasing elevation is related to changes in temperature and water availability (Adriaenssens *et al*., 2017), and in humans, viral richness directly depends on temperature and precipitation globally (Guernier *et al*., 2004). Ultimately, to understand the mechanisms by which viral communities are assembled, it is important to consider both host biology and environmental variability (Bergner *et al*., 2020; Harvell *et al*., 2002; D. Harvell *et al*., 2009; Zamora-Vilchis *et al*., 2012).

The assessment of virus presence and geographic distribution in all components of a lichen symbiosis is an important first step in understanding the potential roles of viruses in lichen ecology and evolution. Future research should address how viral taxa are involved in symbiosis, and specifically explore the effects of variation of viral abundances on host performance across all symbiotic components. For example, why are only some algal strains infected? Here, a follow-up study could explore the potential involvement of the algal-associated *U. phaea* caulimovirus-like 1 in drought tolerance in low elevational populations, where *U. phaea* regularly experiences extended summer droughts (Rolshausen *et al*., 2020). Additionally, future research should include RNA-based approaches, to reveal whether RNA viruses

play a role, and if the putative lichen viruses are active or endogenized viruses (Marzano and Domier, 2016; Nerva *et al*., 2016; Zhang *et al*., 2019).

Conclusion

Our study demonstrates how a long-read metagenomic exploration of the lichen holobiont can be utilized for the reconstruction of novel, complete viral genomes. In addition, tracking viral abundance across host populations provides a widely applicable framework for retrieving meaningful biogeographical determinants of viral distribution. To this end our work provides the first insights into diversity and distribution of viruses that inhabit lichens, suggesting an interplay of host and climate-driven association. This, on the one hand, opens up new research avenues with implications for understanding complex symbioses, and on the other hand, represents a promising source of novel genetic tools for plant biotechnology whose potential remains to date largely unexplored.

Acknowledgements

We thank the Genome Technology Center (RGTC) at Radboudumc for the use of the Sequencing Core Facility (Nijmegen, The Netherlands), which provided the PacBio SMRT sequencing service on the Sequel II platform. We thank Kerry Knudsen and Jana Kocourková (Prague) for their help with collecting the samples, and Tilman Schell (Frankfurt am Main) for his support with bioinformatics.

Author Contributions

D.M., F.D.G. and I.S. conceived the ideas; I.S. collected the data; C.G. and J.O. performed laboratory work; D.M. and F.D.G. analysed data; D.M., F.D.G. and I.S. wrote the manuscript. All authors contributed to the various drafts and gave final approval for publication.

Data Availability Statement

The metagenome and Illumina pooled sequence sequences are available at the Sequence Read Archive under the BioProject PRJNA693984.

References

Adriaenssens, E.M., Kramer, R., van Goethem, M.W., Makhalanyane, T.P., Hogg, I., and Cowan, D.A. (2017) Environmental drivers of viral community composition in Antarctic soils identified by viromics. *Microbiome* **5**: 1–14. https://doi.org/10.1186/s40168-017-0301-7.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Amack, S.C., and Antunes, M.S. (2020) CaMV35S promoter – a plant biology and biotechnology workhorse in the era of synthetic biology. *Curr Plant Biol* **24**: 100179. https://doi.org/10.1016/j.cpb.2020.100179.

Armaleo, D., Müller, O., Lutzoni, F., Andrésson, Ó.S., Blanc, G., Bode, H.B., *et al*. (2019) The lichen symbiosis re-viewed through the genomes of *Cladonia grayi* and its algal partner *Asterochloris glomerata*. *BMC Genomics* **20**: 1–33. https://doi.org/10.1186/s12864-019-5629-x.

Arnold, A.E., Miadlikowska, J., Higgins, K.L., Sarvate, S.D., Gugger, P., Way, A., *et al*. (2009) A phylogenetic estimation of trophic transition networks for ascomycetous fungi: are lichens cradles of symbiotrophic fungal diversification? *Syst Biol* **58**: 283–297. https://doi.org/10.1093/sysbio/syp001.

Asaf, S., Numan, M., Khan, A.L., and Al-Harrasi, A. (2020) *Sphingomonas*: from diversity and genomics to functional role in environmental remediation and plant growth. *Crit Rev Biotechnol* **40**: 138–152. https://doi.org/10.1080/07388551.2019.1709793.

Aschenbrenner, I.A., Cardinale, M., Berg, G., and Grube, M. (2014) Microbial cargo: do bacteria on symbiotic propagules reinforce the microbiome of lichens? *Environ Microbiol* **16**: 3743–3752. https://doi.org/10.1111/1462-2920.12658.

Bak, A., and Emerson, J.B. (2020) Cauliflower mosaic virus (CaMV) biology, management, and relevance to GM plant detection for sustainable organic agriculture. *Front Sustainable Food Syst* **4**: 1–8. https://doi.org/10.3389/fsufs.2020.00021.

Bates, S.T., Nash, T.H., and Garcia-Pichel, F. (2012) Patterns of diversity for fungal assemblages of biological soil crusts from the southwestern United States. *Mycologia* **104**: 353–361. https://doi.org/10.3852/11-232.

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., *et al*. (2013) Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods Ecol Evol* **4**: 914–919. https://doi.org/10.1111/2041-210X.12073.

Bergner, L.M., Orton, R.J., Benavides, J.A., Becker, D.J., Tello, C., Biek, R., and Streicker, D.G. (2020) Demographic and environmental drivers of metagenomic viral diversity in vampire bats. *Mol Ecol* **29**: 26–39. https://doi.org/10.1111/mec.15250.

Boetzer, M., and Pirovano, W. (2014) SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics* **15**: 1–9. https://doi.org/10.1186/1471-2105-15-211.

Breitbart, M., Bonnain, C., Malki, K., and Sawaya, N.A. (2018) Phage puppet masters of the marine microbial realm. *Nat Microbiol* **3**: 754–766. https://doi.org/10.1038/s41564-018-0166-y.

Breitbart, M., Miyake, J.H., and Rohwer, F. (2004) Global distribution of nearly identical phage-encoded DNA

sequences. *FEMS Microbiol Lett* **236**: 249–256. https://doi.org/10.1016/j.femsle.2004.05.042.

Brum, J.R., and Sullivan, M.B. (2015) Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat Rev Microbiol* **13**: 147–159. https://doi.org/10.1038/nrmicro3404.

Buchfink, B., Xie, C., & Huson, D.H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, **12**: 59–60. https://doi.org/10.1038/nmeth.3176

Casano, L.M., Del Campo, E.M., García-Breijo, F.J., Reig-Armiñana, J., Gasulla, F., Del Hoyo, A., *et al*. (2011) Two *Trebouxia* algae with different physiological performances are ever-present in lichen thalli of *Ramalina farinacea*. Coexistence versus competition? *Environ Microbiol* **13**: 806–818. https://doi.org/10.1111/j.1462-2920.2010.02386.x.

Cernava, T., Erlacher, A., Aschenbrenner, I.A., Krug, L., Lassek, C., Riedel, K., *et al*. (2017) Deciphering functional diversification within the lichen microbiota by meta-omics. *Microbiome* **5**: 82. https://doi.org/10.1186/s40168-017-0303-5.

Chow, C.E.T., and Suttle, C.A. (2015) Biogeography of viruses in the sea. *Annu Rev Virol* **2**: 41–66. https://doi.org/10.1146/annurev-virology-031413-085540.

Coutinho, F.H., Silveira, C.B., Gregoracci, G.B., Thompson, C.C., Edwards, R.A., Brussaard, C.P.D., *et al*. (2017) Marine viruses discovered via metagenomics shed light on viral strategies throughout the oceans. *Nat Commun* **8**: 1–12. https://doi.org/10.1038/ncomms15955.

Cubero, O.F., and Crespo, A. (2002) Isolation of nucleic acids from lichens. In *Protocols in Lichenology*, Kranner, I., Beckett, R., and Varma, A. (eds). Berlin, Heidelberg: Springer, pp. 381–391.

Dal Grande, F., Alors, D., Divakar, P.K., Bálint, M., Crespo, A., and Schmitt, I. (2014) Insights into intrathalline genetic diversity of the cosmopolitan lichen symbiotic green alga *Trebouxia decolorans* Ahmadjian using microsatellite markers. *Mol Phylogenet Evol* **72**: 54–60. https://doi.org/10.1016/j.ympev.2013.12.010.

Dal Grande, F., Rolshausen, G., Divakar, P.K., Crespo, A., Otte, J., Schleuning, M., and Schmitt, I. (2018) Environment and host identity structure communities of green algal symbionts in lichens. *New Phytol* **217**: 277–289. https://doi.org/10.1111/nph.14770.

Dal Grande, F., Sharma, R., Meiser, A., Rolshausen, G., Büdel, B., Mishra, B., *et al*. (2017) Adaptive differentiation coincides with local bioclimatic conditions along an elevational cline in populations of a lichen-forming fungus. *BMC Evol Biol* **17**: 1–14. https://doi.org/10.1186/s12862-017-0929-8.

Fernández-Mendoza, F., Fleischhacker, A., Kopun, T., Grube, M., and Muggia, L. (2017) ITS1 metabarcoding highlights low specificity of lichen mycobiomes at a local scale. *Mol Ecol* **26**: 4811–4830. https://doi.org/10.1111/ijlh.12426.

Gallego, V., Sánchez-Porro, C., García, M.T., and Ventosa, A. (2006). Roseomonas aquatica sp. nov., isolated from drinking water. *International Journal of Systematic and Evolutionary Microbiology*, **56**: 2291–2295. https://doi.org/10.1099/ijs.0.64379-0

Garg, N., Zeng, Y., Edlund, A., Melnik, A.V., Sanchez, L.M., Mohimani, H., *et al*. (2016) Microbial Community of a Peltigera. *Msystems* **1**: 1–17. https://doi.org/10.1128/mSystems.00139-16.Editor.

Ghabrial, S.A., Castón, J.R., Jiang, D., Nibert, M.L., and Suzuki, N. (2015) 50-plus years of fungal viruses. *Virology* **479–480**: 356–368. https://doi.org/10.1016/j.virol.2015.02.034.

Göker, M., Scheuner, C., Klenk, H.-P., Stielow, B., and Menzel, W. (2011) Codivergence of Mycoviruses with their hosts. *PLoS One* **6**: e22252. https://doi.org/10.1371/Citation.

Graham, L.E., Trest, M.T., Will-Wolf, S., Miicke, N.S., Atonio, L.M., Piotrowski, M.J., and Knack, J.J. (2018) Microscopic and metagenomic analyses of *Peltigera ponojensis* (Peltigerales, Ascomycota). *Int J Plant Sci* **179**: 241–255. https://doi.org/10.1086/696534.

Greninger, A.L. (2018) A decade of RNA virus metagenomics is (not) enough. *Virus Res* **244**: 218–229. https://doi.org/10.1016/j.virusres.2017.10.014.

Greshake Tzovaras, B., Segers, F.H.I.D., Bicker, A., Dal Grande, F., Otte, J., Anvar, S.Y., *et al*. (2020) What is in *Umbilicaria pustulata*? A metagenomic approach to reconstruct the holo-genome of a lichen. *Genome Biol Evol* **12**: 309–324. https://doi.org/10.1093/gbe/evaa049.

Grube, M., and Berg, G. (2009) Microbial consortia of bacteria and fungi with focus on the lichen symbiosis. *Fungal Biol Rev* **23**: 72–85. https://doi.org/10.1016/j.fbr.2009.10.001.

Grube, M., Cardinale, M., De Castro, J.V., Müller, H., and Berg, G. (2009) Species-specific structural and functional diversity of bacterial communities in lichen symbioses. *ISME J* **3**: 1105–1115. https://doi.org/10.1038/ismej.2009.63.

Grube, M., Cernava, T., Soh, J., Fuchs, S., Aschenbrenner, I., Lassek, C., *et al*. (2015) Exploring functional contexts of symbiotic sustain within lichen-associated bacteria by comparative omics. *ISME J* **9**: 412–424. https://doi.org/10.1038/ismej.2014.138.

Guernier, V., Hochberg, M.E., and Guégan, J.F. (2004) Ecology drives the worldwide distribution of human diseases. *PLoS Biol* **2**: 740–746. https://doi.org/10.1371/journal.pbio.0020141.

Gundlapally, S.R., and Garcia-Pichel, F. (2006). The community and phylogenetic diversity of biological soil crusts in the Colorado Plateau studied by molecular fingerprinting and intensive cultivation. *Microbial Ecology*, **52**: 345–357. https://doi.org/10.1007/s00248-006-9011-6

Harvell, C.D., Mitchell, C.E., Ward, J.R., Altizer, S., Dobson, A.P., Ostfeld, R.S., and Samuel, M.D. (2002) Climate warming and disease risks for terrestrial and marine biota. *Science* **296**: 2158–2162. https://doi.org/10.1126/science.1063699.

Harvell, D., Altizer, S., Cattadori, I.M., Harrington, L., and Weil, E. (2009) Climate change and wildlife diseases: when does the host matter the most? *Ecology* **90**: 912–920. https://doi.org/10.1890/08-0616.1.

Hayes, S., Mahony, J., Nauta, A., and Van Sinderen, D. (2017) Metagenomic approaches to assess bacteriophages in various environmental niches. *Viruses* **9**: 1–22. https://doi.org/10.3390/v9060127.

Hillman, B.I., Annisa, A., and Suzuki, N. (2018) Viruses of plant-interacting fungi. *Adv Virus Res* **100**: 99–116. https://doi.org/10.1016/bs.aivir.2017.10.003.

Hillebrand, H., Blasius, B., Borer, E.T., Chase, J.M., Downing, J.A., Eriksson, B.K., *et al.* (2018). Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring. *Journal of Applied Ecology*, **55**: 169–184. https://doi.org/10.1111/1365-2664.12959

Hodkinson, B.P., Gottel, N.R., Schadt, C.W., and Lutzoni, F. (2012) Photoautotrophic symbiont and geography are major factors affecting highly structured and diverse bacterial communities in the lichen microbiome. *Environ Microbiol* **14**: 147–161. https://doi.org/10.1111/j.1462-2920.2011.02560.x.

Hodkinson, B.P., and Lutzoni, F. (2009) A microbiotic survey of lichen-associated bacteria reveals a new lineage from the Rhizobiales. *Symbiosis* **49**: 163–180. https://doi.org/10.1007/s13199-009-0049-3.

Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007). MEGAN analysis of metagenomic data. *Genome Research*, **17**: 377–386. https://doi.org/10.1101/gr.5969107

Jiang, C.Y., Dai, X., Wang, B.J., Zhou, Y.G., and Liu, S.J. (2006) Roseomonas lacus sp. nov., isolated from freshwater lake sediment. *International Journal of Systematic and Evolutionary Microbiology*, **56**: 25–28. https://doi.org/10.1099/ijs.0.63938-0

Jiang, D.F., Wang, H.Y., Si, H.L., Zhao, L., Liu, C.P., and Zhang, H. (2017) Isolation and culture of lichen bacteriobionts. *Lichenologist* **49**: 175–181. https://doi.org/10.1017/S0024282917000044.

Kappen, L. (2000) Some aspects of the great success of lichens in Antarctica. *Antarct Sci* **12**: 314–324. https://doi.org/10.1017/s0954102000000377.

Khan, A.L., Waqas, M., Asaf, S., Kamran, M., Shahzad, R., Bilal, S., *et al*. (2017) Plant growth-promoting endophyte *Sphingomonas* sp. LK11 alleviates salinity stress in *Solanum pimpinellifolium*. *Environ Exp Bot* **133**: 58–69. https://doi.org/10.1016/j.envexpbot.2016.09.009.

Khan, A.L., Waqas, M., Kang, S.M., Al-Harrasi, A., Hussain, J., Al-Rawahi, A., *et al*. (2014) Bacterial endophyte *Sphingomonas* sp. LK11 produces gibberellins and IAA and promotes tomato plant growth. *J Microbiol* **52**: 689–695. https://doi.org/10.1007/s12275-014-4002-7.

Kieft, K., Zhou, Z., and Anantharaman, K. (2020) VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome* **8**: 1–23. https://doi.org/10.1186/s40168-020-00867-0.

Kim, H., Nishiyama, M., Kunito, T., Senoo, K., Kawahara, K., Murakami, K., and Oyaizu, H. (1998) High population of *Sphingomonas* species on plant surface. *J Appl Microbiol* **85**: 731–736. https://doi.org/10.1111/j.1365-2672.1998.00586.x.

Kolmogorov, M., Rayko, M., Yuan, J., Polevikov, E., and Pevzner, P. (2019) metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods* **17**: 1103–1110. https://doi.org/10.1101/637637.

Kotta-Loizou, I., and Coutts, R.H.A. (2017) Mycoviruses in Aspergilli: a comprehensive review. *Front Microbiol* **8**: 1699. https://doi.org/10.3389/fmicb.2017.01699.

Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with bowtie 2. *Nat Methods* **9**: 357–359. https://doi.org/10.1038/nmeth.1923.

Lavigne, R., Molineux, I. J., and Kropinski, A. M. (2011) Virus taxonomy: ninth report of the international committee on taxonomy of viruses. *ICTV 9th Report*.

Luis, M., Márquez, R.S.R., *et al*. (2007) A virus in a fungus in a plant. *Science* **315**: 513–516.

Luo, Y., Wang, F., Huang, Y., Zhou, M., Gao, J., Yan, T., *et al*. (2019) Sphingomonas sp. Cra20 increases plant growth rate and alters rhizosphere microbial community structure of *Arabidopsis thaliana* under drought stress. *Front Microbiol* **10**: 1221. https://doi.org/10.3389/fmicb.2019.01221.

Manrique, P., Bolduc, B., Walk, S.T., Van, O., van der Oost, J., De Vos, W.M., and Young, M.J. (2016) Healthy human gut phageome. *Proc Natl Acad Sci U S A* **113**: 10400–10405. https://doi.org/10.1073/pnas.1601060113.

Marzano, S.Y.L., and Domier, L.L. (2016) Novel mycoviruses discovered from metatranscriptomics survey of soybean phyllosphere phytobiomes. *Virus Res* **213**: 332–342. https://doi.org/10.1016/j.virusres.2015.11.002.

Massart, S., Olmos, A., Jijakli, H., and Candresse, T. (2014) Current impact and future directions of high throughput sequencing in plant virus diagnostics. *Virus Res* **188**: 90–96. https://doi.org/10.1016/j.virusres.2014.03.029.

Mayjonade, B., Gouzy, J., Donnadieu, C., Pouilly, N., Marande, W., Callot, C., *et al*. (2016) Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. *Biotechniques* **61**: 203–205. https://doi.org/10.2144/000114460.

McGrath, S., Fitzgerald, G.F., and Van Sinderen, D. (2004) The impact of bacteriophage genomics. *Curr Opin Biotechnol* **15**: 94–99. https://doi.org/10.1016/j.copbio.2004.01.007.

Menzel, P., Ng, K.L., and Krogh, A. (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* **7**: 11257. https://doi.org/10.1038/ncomms11257.

Moran, N.A., Degnan, P.H., Santos, S.R., Dunbar, H.E., and Ochman, H. (2005) The players in a mutualistic symbiosis: insects, bacteria, viruses, and virulence genes. *Proc Natl Acad Sci U S A* **102**: 16919–16926. https://doi.org/10.1073/pnas.0507029102.

Moya, A., Elena, S.F., Bracho, A., Miralles, R., and Barrio, E. (2000) The evolution of RNA viruses: a population genetics view. *Proc Natl Acad Sci U S A* **97**: 6967–6973. https://doi.org/10.1073/pnas.97.13.6967.

Nakayama, N., Tsuge, T., Asakawa, S., and Kimura, M. (2009) Morphology, host range and phylogenetic diversity of *Sphingomonas* phages in the floodwater of a Japanese paddy field. *Soil Sci Plant Nutr* **55**: 53–64. https://doi.org/10.1111/j.1747-0765.2008.00332.x.

Nayfach, S., Camargo, A.P., Eloe-Fadrosh, E., Roux, S., and Kyrpides, N. (2020) CheckV: assessing the quality of metagenome-assembled viral genomes. *Nat Biotechnol* **39**: 578–585. https://doi.org/10.1101/2020.05.06.081778.

Nerva, L., Ciuffo, M., Vallino, M., Margaria, P., Varese, G.C., Gnavi, G., and Turina, M. (2016) Multiple approaches for the detection and characterization of viral and plasmid

symbionts from a collection of marine fungi. *Virus Res* **219**: 22–38. https://doi.org/10.1016/j.virusres.2015.10.028.

Noh, H.J., Lee, Y.M., Park, C.H., Lee, H.K., Cho, J.C., and Hong, S.G. (2020) Microbiome in *Cladonia squamosa* is vertically stratified according to microclimatic conditions. *Front Microbiol* **11**: 1–13. https://doi.org/10.3389/fmicb.2020.00268.

Noh, H.J., Park, Y., Hong, S.G., and Lee, Y.M. (2021) Diversity and physiological characteristics of antarctic lichens-associated bacteria. *Microorganisms* **9**: 1–15. https://doi.org/10.3390/microorganisms9030607.

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017) MetaSPAdes: a new versatile metagenomic assembler. *Genome Res* **27**: 824–834. https://doi.org/10.1101/gr.213959.116.

Nuss, D.L. (2005) Hypovirulence: mycoviruses at the fungal-plant interface. *Nat Rev Microbiol* **3**: 632–642. https://doi.org/10.1038/nrmicro1206.

Paez-Espino, D., Eloe-Fadrosh, E.A., Pavlopoulos, G.A., Thomas, A.D., Huntemann, M., Mikhailova, N., *et al*. (2016) Uncovering Earth's virome. *Nature* **536**: 425–430. https://doi.org/10.1038/nature19094.

Parker, B.J., and Brisson, J.A. (2019) A laterally transferred viral gene modifies aphid wing plasticity. *Curr Biol* **29**: 2098–2103.e5. https://doi.org/10.1016/j.cub.2019.05.041.

Paul, F., Otte, J., Schmitt, I., and Dal Grande, F. (2018) Comparing sanger sequencing and high-throughput metabarcoding for inferring photobiont diversity in lichens. *Sci Rep* **8**: 1–7. https://doi.org/10.1038/s41598-018-26947-8.

Peksa, O., and Škaloud, P. (2011) Do photobionts influence the ecology of lichens? A case study of environmental preferences in symbiotic green alga *Asterochloris* (Trebouxiophyceae). *Mol Ecol* **20**: 3936–3948. https://doi.org/10.1111/j.1365-294X.2011.05168.x.

Petrzik, K., Koloniuk, I., Sehadová, H., and Sarkisova, T. (2019) Chrysoviruses inhabited symbiotic fungi of lichens. *Viruses* **11**: 1–15. https://doi.org/10.3390/v11121120.

Petrzik, K., Vondrák, J., Barták, M., Peksa, O., and Kubešová, O. (2014) Lichens-a new source or yet unknown host of herbaceous plant viruses? *Eur J Plant Pathol* **138**: 549–559. https://doi.org/10.1007/s10658-013-0246-z.

Petrzik, K., Vondrák, J., Kvíderová, J., and Lukavský, J. (2015) Platinum anniversary: virus and lichen alga together more than 70 years. *PLoS One* **10**: 1–12. https://doi.org/10.1371/journal.pone.0120768.

Prasanna, H.C., Sinha, D.P., Verma, A., Singh, M., Singh, B., Rai, M., and Martin, D.P. (2010) The population genomics of begomoviruses: global scale population structure and gene flow. *Virol J* **7**: 1–12. https://doi.org/10.1186/1743-422X-7-220.

Quinlan, A.R., and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842. https://doi.org/10.1093/bioinformatics/btq033.

R Core Team. (2019) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ramírez-Fernández, L., Zúñiga, C., Carú, M., and Orlando, J. (2014) Environmental context shapes the bacterial community structure associated to *Peltigera* cyanolichens growing in Tierra del Fuego, Chile. *World J Microbiol Biotechnol* **30**: 1141–1144. https://doi.org/10.1007/s11274-013-1533-8.

Remnant, E.J., Shi, M., Buchmann, G., Blacquière, T., Holmes, E.C., Beekman, M., and Ashe, A. (2017) A diverse range of novel RNA viruses in geographically distinct honey bee populations. *J Virol* **91**: 1–19.

Rolshausen, G., Dal Grande, F., Sadowska-Deś, A.D., Otte, J., and Schmitt, I. (2018) Quantifying the climatic niche of symbiont partners in a lichen symbiosis indicates mutualist-mediated niche expansions. *Ecography* **41**: 1380–1392. https://doi.org/10.1111/ecog.03457.

Rolshausen, G., Hallman, U., Grande, F.D., Otte, J., Knudsen, K., and Schmitt, I. (2020) Expanding the mutualistic niche: parallel symbiont turnover along climatic gradients. *Proc R Soc B: Biol Sci* **287**: 20192311. https://doi.org/10.1098/rspb.2019.2311.

Roossinck, M.J. (2015a) Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles. *Front Microbiol* **5**: 767. https://doi.org/10.3389/fmicb.2014.00767.

Roossinck, M.J. (2015b) Move over, bacteria! viruses make their mark as mutualistic microbial symbionts. *J Virol* **89**: 6532–6535. https://doi.org/10.1128/jvi.02974-14.

Roux, S., Adriaenssens, E.M., Dutilh, B.E., Koonin, E.V., Kropinski, A.M., Krupovic, M., *et al*. (2019) Minimum information about an uncultivated virus genome (MIUVIG). *Nat Biotechnol* **37**: 29–37. https://doi.org/10.1038/nbt.4306.

Roux, S., Páez-Espino, D., Chen, I.M.A., Palaniappan, K., Ratner, A., Chu, K., *et al.* (2021). MG/VR v3: An integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. *Nucleic Acids Research*, **49**: 764–775. https://doi.org/10.1093/nar/gkaa946

Shi, M., Neville, P., Nicholson, J., Eden, J.S., Imrie, A., and Holmes, E.C. (2017) High-resolution metatranscriptomics reveals the ecological dynamics of mosquito-associated RNA viruses in Western Australia. *J Virol* **91**: 1–17.

Short, S.M. (2012) The ecology of viruses that infect eukaryotic algae. *Environ Microbiol* **14**: 2253–2271. https://doi.org/10.1111/j.1462-2920.2012.02706.x.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210–3212. https://doi.org/10.1093/bioinformatics/btv351.

Simmonds, P., Adams, M.J., Benk, M., Breitbart, M., Brister, J.R., Carstens, E.B., *et al*. (2017) Consensus statement: virus taxonomy in the age of metagenomics. *Nat Rev Microbiol* **15**: 161–168. https://doi.org/10.1038/nrmicro.2016.177.

Singh, G., Dal Grande, F., Divakar, P.K., Otte, J., Crespo, A., and Schmitt, I. (2017) Fungal–algal association patterns in lichen symbiosis linked to macroclimate. *New Phytol* **214**: 317–329. https://doi.org/10.1111/nph.14366.

Sommers, P., Chatterjee, A., Varsani, A., and Trubl, G. (2021) Integrating viral metagenomics into an ecological framework. *Annu Rev Virol* **8**: 133–158. https://doi.org/10.1146/annurev-virology-010421-053015.

Suttle, C. (2005) Viruses in the sea. *Nature* **437**: 356–361.
Vargas Castillo, R., and Beck, A. (2012) Photobiont selectivity and specificity in *Caloplaca* species in a fog-induced community in the Atacama Desert, northern Chile. *Fungal Biol* **116**: 665–676. https://doi.org/10.1016/j.funbio.2012.04.001.
Werth, S., and Sork, V.L. (2014) Ecological specialization in *Trebouxia* (Trebouxiophyceae) photobionts of *Ramalina menziesii* (Ramalinaceae) across six range-covering ecoregions of western North America. *Am J Bot* **101**: 1127–1140. https://doi.org/10.3732/ajb.1400025.
Wommack, K.E., Ravel, J., Hill, R.T., Chun, J., and Colwell, R.R. (1999) Population dynamics of Chesapeake bay virioplankton: total-community analysis by pulsed-field gel electrophoresis. *Appl Environ Microbiol* **65**: 231–240. https://doi.org/10.1128/aem.65.1.231-240.1999.
Wu, Y., Wang, Y., Li, J., Li, W., Zhang, L., Li, Y., *et al*. (2014) Development of a general method for detection and quantification of the P35S promoter based on assessment of existing methods. *Sci Rep* **4**: 1–13. https://doi.org/10.1038/srep07358.
Zamora-Vilchis, I., Williams, S.E., and Johnson, C.N. (2012) Environmental temperature affects prevalence of blood parasites of birds on an elevation gradient: implications for disease in a warming climate. *PLoS One* **7**: e39208. https://doi.org/10.1371/journal.pone.0039208.
Zhang, Y.Z., Chen, Y.M., Wang, W., Qin, X.C., and Holmes, E.C. (2019) Expanding the RNA virosphere by unbiased metagenomics. *Annu Rev Virol* **6**: 119–139. https://doi.org/10.1146/annurev-virology-092818-015851.
Zhang, Y.Q., Yu, L.Y., Wang, D., Liu, H.Y., Sun, C.H., Jiang, W., *et al.* (2008). Roseomonas vinacea sp. nov., a Gram-negative coccobacillus isolated from a soil sample. *International Journal of Systematic and Evolutionary Microbiology*, **58**: 2070–2074. https://doi.org/10.1099/ijs.0.65789-0

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Pool sequenced *U. phaea* populations and the number of mapped reads against viral genomes and their putative hosts.

Table S2. Mapping results of photobiont haplotype transcriptomes against *U. phaea* caulimovirus-like 1 and S3586 genome.

Table S3. Annotation of viral genes as implemented in VIBRANT.

Table S4. Mean coverage of putative phage hosts.

Fig. S1. The abundance of phage 3 (a, *U. phaea* siphovirus-like 1) and phage 4 (b, *U. phaea* siphovirus-like 1) across the elevational gradients. Linear models of read abundances were fitted. Raw data are shown in red (Mt. San Jacinto) and blue (Sierra Nevada) circles. Dashed lines show a non-significant effect of elevation on viral abundance ($p > 0.05$). Viruses were present in all populations, except viral scaffold1070, which was not present in four of eleven populations (Uph23, Uph25, Uph27, Uph28; Table S1).

Fig. S2. Elevational distribution of the *U. phaea* caulimovirus-like 1 and the two dominant algal haplotypes reconstructed from the pool-seq data.

Fig. S3. Relative abundance of *U. phaea* caulimovirus-like 1 (S3586) RNA in algal cultures.