## Title

# Viruses help rewire carbon metabolism in a methane-suppressed peat microcosm

## Running Head

Viruses rewire microcosm carbon metabolism

## Authors

James Riddell V[1,2] (https://orcid.org/0000-0001-5052-4513), Rokaiya Nurani Shatadru[1,2] (https://orcid.org/0000-0002-9559-9583), Garrett J Smith[2,3] (https://orcid.org/0000-0002-1014-5102), Bridget B McGivern[2,4] (https://orcid.org/0000-0001-9023-0018), Jared B Ellenbogen[2,5] (https://orcid.org/0000-0003-1692-9044), Sophie K Jurgensen[2,5] (https://orcid.org/0000-0001-5399-510X), Ami Fofana[1,2], Malak M Tfaily[2,6] (https://orcid.org/0000-0002-3036-2833), Kelly C Wrighton[2,5] (https://orcid.org/0000-0003-0434-4217) & Matthew B Sullivan[1,2,3,7] (https://orcid.org/0000-0001-8398-8234, sullivan.948@osu.edu)

## Affiliations

1. Department of Microbiology, The Ohio State University, Columbus 43210, OH, USA
2. EMERGE Biology Integration Institute, Columbus, OH, USA
3. Center of Microbiome Science, The Ohio State University, Columbus 43210, OH, USA
4. Department of Chemistry and Biochemistry, University of Wisconsin-Eau Claire, Eau Claire 54703, WI, USA
5. Department of Soil and Crop Sciences, Colorado State University, Fort Collins 80523, CO, USA
6. Department of Environmental Science, University of Arizona, Tucson 85721, AZ, USA
7. Department of Civil, Environmental and Geodetic Engineering, Ohio State University, Columbus, OH, USA

## Abstract

Arctic permafrosts are rapidly thawing in response to climate change, stimulating microbial activity and release of additional greenhouse gases, such as methane. Recently, catechin amendment of thawed permafrost soil microcosms demonstrated >80% decrease in methane production over 35 days compared to unamended controls, with metagenome, metatranscriptome and metabolome analyses revealing a shift in prokaryotic carbon metabolism from a syntrophic network feeding methanogens to one dominated by catechin fermentation. Here we leverage these same data to investigate potential virus impacts on this microbial community metabolism shift. In total, 900 DNA virus operational taxonomic units (vOTUs) were identified as actively lytic based on transcribed structural and lysis genes. Of these, 41% were predicted to infect at least one of 56 transcriptionally active prokaryote genera representing 13 phyla. The most active vOTUs were predicted to infect key catechin-degrading genera including *Clostridium* and undescribed *Bacillota* genus JAGFXR01. Temporally, viral communities responded later than prokaryotes to catechin amendment, reflecting a virus production lag to targeting key microbial responders. A single induced vOTU predicted to infect JAGFXR01 dominated the catechin-amended viral community response, with relative abundances 20-156 times higher than its host suggesting intense viral lysis that could release degraded catechin intermediates. Indeed, examination of catechin intermediate gene expression in non-catechin-degraders revealed that they were elevated in catechin-amended samples, suggesting JAGFXR01 lysis products were taken up by non-catechin-degraders as part of community carbon metabolism rewiring. Together, these results place viruses at the heart of carbon rewiring that modulates ecosystem outputs of climate-critical thawing permafrosts.


## Introduction

Microbes form partnerships that are central to nearly all life on Earth, yet anthropogenic climate change is fundamentally altering how these communities function [1]. Warning signs are emerging across the planet: climate-intensified extreme weather events are shifting microbial disease ranges [2], altering coral-microbe symbioses [3], and destabilizing plant-microbe interactions [4–6]. In some systems, climate change may trigger microbial feedback loops that accelerate warming itself. Arctic permafrost thaw exemplifies this risk—as temperatures rise, microbial decomposition of millennia-old soil organic matter releases substantial carbon to the atmosphere [7–9]. At Stordalen Mire (Abisko, Sweden), a long-term, extensively studied subarctic peatland, permafrost thaw over the past 50 years has doubled the extent of fully-thawed fens [10], which emit 4.5 times more methane than intact palsas and partially frozen bogs, driven by increased microbial methanogenesis [11, 12]. As the climate crisis intensifies and microbiome engineering emerges as a potential mitigation strategy [13, 14], identifying interventions to suppress methane production without disrupting fundamental soil microbiome functions has become an urgent priority.

Towards this, recent work at Stordalen Mire has demonstrated catechin, a common polyphenol in Mire flora [15, 16], can decrease methane emissions by >80% in thawed fen peat microcosms [17] (**Fig. 1A**). Metagenomics, metatranscriptomics, and metabolomics data collected from these microcosms revealed catechin amendments stimulated community-wide metabolic rewiring where dominant heterotrophic carbohydrate degraders declined, and

catechin-degrading taxa, including undescribed *Bacillota* genus JAGFXR01 and *Clostridium* spp., became central to carbon decomposition. Because these newly abundant and active catechin degraders competitively consumed hydrogen and secreted acidic by-products, they collectively suppressed hydrogenotrophic methanogenesis, which is the dominant methanogenesis pathway in these fens [17]. Given such drastic rewiring of prokaryotic processes, we wondered whether viruses could participate in, amplify, and/or additionally impact this metabolic rewiring. Viruses are already known to be ubiquitous and play an active role across diverse ecosystems, contributing to host mortality and lysis-mediated nutrient release [18] and altering the metabolic objective of host cells during infection (termed virocells) [19]. In soils, there is also mounting evidence viruses influence carbon cycling as inferred from experimental addition of viruses to soils and virus-encoded auxiliary metabolic genes involved in carbon cycling, and linkage of virus dynamics to key carbon cycling host species [20–24]. The viral response to rapid environmental change, such as abrupt thaw and substrate amendment, and their impact on subsequent community-wide metabolic shifts remains virtually unknown. This knowledge gap represents an opportunity to establish baseline data needed to integrate viruses into multi-scale models that predict the fate of carbon in ecosystems worldwide. [25].

Here, we investigate viral community dynamics and potential functional roles during catechin-mediated methane suppression in permafrost peat microcosms using the same multi-omics dataset from the earlier fen microcosm prokaryote-focused work [17]. Specifically, we leveraged viral identification [26] and host prediction tools [27] to: (i) identify actively replicating viruses and their predicted prokaryotic hosts, (ii) quantify viral community responses to catechin amendment across time, and (iii) evaluate whether viral lysis of key catechin degraders could modulate carbon flow by releasing partially metabolized intermediates of catechin metabolism into the broader microbial community through a viral shunt mechanism [28]. Together, these findings suggest that viruses may play a role in redistributing catechin-derived carbon intermediates, linking specialized degraders with the broader microbial community leading to net methane suppression.

# Materials and Methods

## Sample Collection

See *Peat sampling*, *Incubation set-up,* and *Nucleic acids extraction* in [17] for full description. Briefly, a total of 40 anoxic microcosms were constructed from peat from a fen in Stordalen Mire (Abisko, Sweden) 9-19cm below the surface, where half were left as controls and half were catechin-amended (**Fig. 1A**). To generate bulk metagenomes, one microcosm per treatment (catechin, unamended) was sacrificially sampled at days 0, 14, 21, and 35 post-amendment. DNA was extracted and used for Illumina sequencing library construction, which was sequenced up to 500M reads with a median depth of 421M reads. To generate bulk prokaryote metatranscriptome sequencing data, and assess variation, three microcosms per amendment were sacrificed at days 0, 7, 14, 21, and 35 post-amendment. RNA was extracted, rRNA depleted, and DNase treated. The resulting RNA was reverse-transcribed to cDNA, and Illumina sequencing libraries were constructed and sequenced to a depth of up to 100 million reads (median depth: 99 million reads). Day 0 only had metagenomic and metatranscriptomic sequencing for unamended samples because it was sacrificed immediately after the

microcosms were sealed and served as the initial timepoint for both treatments. Sequencing sample accession numbers are available in Supplementary Data 1.

## Metagenome Assembly and MAG recovery and taxonomic/functional/transcriptomic assessment

See *Metagenome sequencing, assembly, and binning* in [17] for full description. Briefly, metagenome sequencing libraries were assembled with three strategies: (1) Samples individually assembled with MEGAHIT (v1.2.9) [28], (2) Iterative assembly of select individual samples, whereby trimmed metagenome reads from individual samples were mapped against medium and high-quality MAGs generated from the individual assemblies, and reads that did not map were assembled again. Assembled contigs were binned, and the process was repeated until no or few medium or high-quality MAGs were recovered. (3) Co-assembly of different samples with MEGAHIT, where assembled contigs were binned using the combined sample reads to generate coverage profiles. MAGs resulting from the microcosm-derived metagenomes were combined with a set of MAGs that were recovered from a decade of sampling at Stordalen Mire. This full MAG set was dereplicated at 99% ANI to produce approximately 2,302 species-level ecological units.

MAGs were then analysed for taxonomy, functional content, and activity as follows. First, MAG taxonomy was inferred using the Genome Taxonomy Database Toolkit (GTDB-tk v2.3.0 r214) [29]. Second, MAGs were annotated using DRAM (v1.4.4) [30] and CAMPER (v1) [31]. Third, relative expression was approximated with gene length corrected trimmed mean of M-values (GeTMM) [32]. Relative expression of MAG genes were taken from extended data computed previously [17]. MAG relative expression was computed as the average relative expression per gene, calculated by summing the GeTMM values of active genes and dividing by the total number of genes per MAG.

## Virus database construction

Viral contigs were mined from (i) all assembled contigs >5kb from Stordalen peat fen microcosm metagenomic samples (https://zenodo.org/records/17652154), (ii) medium and high-quality (est. completeness >50%, <10% contamination) MAGs recovered from the microcosms [17], and (iii) vOTUs previously identified from Stordalen Mire samples [22, 33]. Viral contigs were identified with GeNomad (v1.8.0) [26] using default settings. GeNomad also trimmed host regions from contigs it identified as provirus. Next, contigs considered viral based on GeNomad default post-classification filtering were further quality scored with CheckV (v1.0.3) [34] using default settings, which identified additional provirus sequences, trimmed host regions, estimated viral contig completeness, and estimated contamination from non-viral genes. Viral contigs were then clustered into viral operational taxonomic units (vOTUs) in accordance with the Minimum Information for an Uncultivated Virus Genome (MIUViG) standards [35]. vOTUs were annotated with prodigal-gv (v2.6.3) [36] via the GeNomad annotate module, and inputs for DRAMv functional annotation were prepared with the MVP pipeline (v1.0) [19] module 06. DRAMv (v1.4) [30] was then used to functionally annotate vOTUs.

As an additional quality control step, 100 vOTUs were randomly chosen for manual inspection of functional annotations (Supplementary Data 2). Out of the 100 vOTUs, six were manually classified as “not viral” due to either 1) possessing a transposase but no other viral-

like genes or 2) having no viral genes detected by CheckV and >30% of genes annotated as non-viral genes by DRAMv. Based on these rules, we filtered out 470 vOTUs (Supplementary Data 3) resulting in 9,092 vOTUs (Supplementary Data 4). All 9,092 vOTUs were taxonomically classified with vConTACT3 (v3.0.5) [38] with virus reference database v228 using default parameters (Supplementary Data 5).

## Classifying active vOTUs

See *Metatranscriptome Sequencing and Analysis* in [17] for sequencing, read quality filtering, and subsampling procedures. To determine which vOTUs were transcriptionally active, metatranscriptomes were randomly subsampled to 50M pairs of trimmed reads and then aligned to the vOTU reference database using Bowtie2 (v2.5.1) [39] with the following parameters: -D 10 -R 2 -N 1 -L 22 -i S,0,2.50. Output SAM files were converted to sorted BAM files using samtools (v1.18) [40]. Reads mapping <90% average sequence identity to vOTUs were filtered out of the sorted BAM files using bbtools (v38.97) reformat.sh [41] with the parameters: idfilter=0.90 pairedonly=t primaryonly=t. The 90% cutoff was chosen based on viral population read mapping cutoffs established previously [42]. The number of reads mapped to each vOTU gene were counted by first predicting open reading frames with prodigal-gv (v2.6.3), then using htseq-count (v.0.13.5) with the following parameters: -a 0 -t CDS -i ID --stranded=reverse, which filtered out any reads that ambiguously mapped to more than one gene. Read counts <5 were set to zero, as in the previous study [17] and another soil virus metatranscriptomics study [43]. vOTU genes were annotated with DRAMv [30] and GeNomad [26], and database annotations were concatenated into a single string for each gene.

We considered a vOTU transcriptionally active in a sample if it had (i) ≥5 reads mapping to ≥1 gene that contained any of the substrings “capsid”, “tail”, “package”, “packaging”, “portal”, “lysis”, “terminase”, “spike”, “baseplate”, “internal virion”, “neck”, or “prohead” in its concatenated annotation, a similar strategy to [44] for classifying active-lytic viruses, and (ii) ≥10% of the bases on the viral contig were covered by at least one read using bedtools genomecov [45]. If a vOTU did not meet the requirements of (i) and (ii) in a sample, its coverage was set to zero in that sample.

## vOTU characterizations and analyses

**Activity:** The relative expression of vOTUs in metatranscriptomes was calculated using GeTMM normalization method [32] from the EdgeR package (v4.4.2) in R (v4.4). Read counts per vOTU, per sample, that met the criterion above (not set to zero) were GeTMM normalized, then the GeTMM values of each vOTU’s active genes were summed and divided by the total number of genes within the vOTU per sample, yielding an average vOTU GeTMM expression per sample. To determine the rank abundance of vOTUs under different treatments, we identified the maximum GeTMM relative expression attained by each vOTU across all replicate samples within both the catechin-amended and unamended treatments. Species accumulation (collector’s curve), rank abundance, and visualizations were done in Python (v3.12).

**Alpha and Beta Diversity:** vOTU Shannon diversity was calculated from active vOTU GeTMM values rounded up to the nearest whole number using scikit-bio (v0.6.3) [46] in Python (v3.12). The same method was used to determine MAG Shannon diversity. vOTU beta diversity was determined using Hellinger-transformed Bray-Curtis dissimilarity of vOTU GeTMM values

using vegan (v2.6-10) [47] in R. Beta diversity was visualized using principal component analysis. Samples were clustered with affinity propagation [48] based on vOTU GeTMM values. The same approach was used to calculate and visualize MAG beta diversity.

**Virus-host prediction:** Virus-host predictions were made using iPHoP (v1.3.2) [27]. A custom iPHoP host database was constructed by adding the 2,302 MAGs identified from microcosm samples and previous Stordalen analyses to the iPHoP standard database (iPHoP_db_Aug23_rw release). iPHoP was run on vOTU sequences using the custom database and default settings, with vOTU-MAG predictions available in Supplementary Data 6. vOTU-MAG predictions were visualized with Cytoscape [49], and metadata from the node table available in Supplementary Data 7.

**Auxiliary metabolic gene (AMG) identification:** AMGs were identified from active vOTUs using DRAMv annotations. A gene was considered an auxiliary metabolic gene if it contained an “M” (metabolic) flag, but no “B” (consecutive metabolic genes) or “F” (near-end-of-contig) flags, and was both upstream and downstream of a gene with a “V” (viral) flag [30, 50, 51]. All detected AMG annotations are available in Supplementary Data 8.

**Indicator Species Analysis:** Indicator values were computed per vOTU using GeTMM values. The five sample clusters formed by affinity propagation were used as the sample groups to compare vOTU relative abundance between. Indicator values were computed using the multipatt function from the indicspecies package in R with parameter duleg=False to consider combinations of site groups [22]. vOTUs with a p-value < 0.01 were classified as indicators (Supplementary Data 9).

**Metatranscriptomic relative abundance correlation between indicator vOTUs and their predicted host MAGs:** The correlation between indicator vOTUs and their predicted host relative expression values across all metatranscriptome samples was calculated using the pearsonr function from the scipy.stats library for JAGFXR01, *Clostridium*, *Pseudomonas_E*, and *Paludibacter* infection networks predicted by iPHoP. Correlations were considered significant if the Benjamini-Hochberg false discovery rate (adjusted *P* value) was <0.05.

**Metagenomic relative abundance comparisons between viruses and their predicted hosts:** Metagenomic samples were mapped back to vOTUs and MAGs with Bowtie2 (v2.5.1) [39] with the following parameters: -D 10 -R 2 -N 1 -L 22 -i S,0,2.50. Output SAM files were converted to sorted BAM files using samtools [40]. Relative abundance was determined using CoverM [53] with the following flags for vOTUs: mode=trimmed_mean, min-read-percent-identity=90, min-alignment-fraction=75, min-covered-fraction=10, and for MAGs: mode=mean, min-read-percent-identity=97, min-alignment-fraction=75, min-covered-fraction=75.

**Comparative genomics and transcriptome coverage analysis:** Active vOTUs predicted to infect MAGs in the genus JAGFXR01 were clustered in VirClust with default settings [54]. Briefly, open reading frames were predicted using MetaGeneAnnotator [55], and compared using all-vs-all BLASTP. Then, genome distances were calculated using markov clustering. Read pileup was calculated for contig_591846 in catechin-amended samples using samtools mpileup based on the filtered, sorted BAM files generated from metatranscriptomic samples. Contig_591846 genes were visualized with gggenomes [56], and read pileups were visualized with ggridges [57] in R.

# Results + Discussion

## Experimental design, dataset overview, and virus recovery

Our prior work leveraged climate-sensitive peat microcosms from a Stordalen Mire fen to evaluate catechin-suppressed methane production from the prokaryotic community perspective [17]. Catechin amendment decreased methane emissions by >80% from the fen microcosms. To uncover the mechanism/s underlying this inhibition, bulk metagenomes, metatranscriptomes, and metabolomes were obtained from the microcosms along a 35 day time course (**Fig. 1A**, see Methods for details) and analyzed for metabolic signal. These past MAG-resolved metatranscriptome data analyses revealed prokaryotic community expression was dominated by *Clostridium* and *Pseudomonas_E* in nearly all samples, with *Paludibacter* enriched in unamended samples and undescribed *Bacillota* genus JAGFXR01 enriched in catechin-amended samples (**Fig. 1B**). These dominant taxa played key roles in unamended and catechin-amended carbon metabolism (**Fig. 1C**), namely polysaccharide metabolism and catechin and phenolic acid metabolism, respectively. Together these findings revealed a microbial community-wide metabolic transition from a syntrophic network that feeds methanogenesis to a catechin fermentation network. Here we evaluate if, and to what extent, viruses play a role in this community-wide metabolic transition by identifying actively lytic viruses, predicting their hosts, and investigating how viral communities responded to catechin amendment in the context of diverse biomolecular measurements.

We first established relevant reference genome resources by compiling a total of 9,092 DNA viral operational taxonomic units (vOTUs) identified from previous work in Stordalen Mire [22, 33] and current analyses of the microcosm metagenome assemblies (this study). We then were able to classify ~10% (900 of the 9,092 vOTUs) as active in at least one microcosm metatranscriptome (**Fig. S1A,** see methods). Because no established standard exists for defining viral activity from metatranscriptome data [58], we employed relatively conservative activity thresholds requiring (i) ≥5 reads mapped to a gene involved in viral particle formation or lysis and (ii) >10% of bases on the viral contig were covered by at least one read. These thresholds ensured the viruses classified as active were transcribing genes associated with viral particle formation or lysis, a hallmark of lytic activity [59], and decreased false positive activity classification from non-specific read mapping by requiring a minimum coverage breadth [60]. Sensitivity analyses classifying activity more permissively as in other studies (e.g., one read mapped per gene per 10kb [22] or >10% coverage [44] thresholds) captured similar ecological patterns (**Fig. S2A**). Active vOTUs were relatively well sampled in the metatranscriptomes (**Fig. S1C**), and even lowly active vOTUs were captured based on rank abundance (**Fig. S1D**).

## Viral community activity responses differ across treatments

We evaluated how community-expressed alpha diversity (see methods) of these active vOTUs differed between treatments and/or over time (**Fig. 2A**). Both treatments revealed vOTU alpha diversity changed minimally by day 7, declined at day 14, rebounded at day 21, and declined again at day 35. In catechin-amended samples, this pattern was similar but more pronounced, and overall alpha diversity was lower compared to unamended samples. Catechin-amended sample vOTU alpha diversity decreased sharply from day 7 to 14 coincident with a four-fold increase in the proportion of all reads mapping to active vOTUs (from 0.25 to 1%),

284 whereas unamended samples maintained higher diversity and the proportion of reads mapping
285 to vOTUs remained low. Together, these results suggest a catechin-specific induction of lytic
286 activity in a subset of vOTUs.
287 We investigated beta diversity to determine how similar viral community expression was
288 among samples. This revealed viral community expression was similar at days 0 and 7, then
289 diverged between treatments by day 14 and remained distinct thereafter (**Fig. 2B**). This
290 suggests a different virus community was enriched in catechin-amended day 14-35 samples
291 compared to day 7 and unamended day 14-35 samples. Additionally, community-level
292 ecological patterns still emerged despite high turnover in active vOTUs between samples,
293 where the average vOTU was active in only two samples (**Fig. S1E**), which was similar to
294 previous observations in soil virus environmental studies [33, 61, 62].
295 We wondered how these viral community dynamic patterns compared to those derived
296 from the MAGs captured in these data. These analyses revealed that expression-based MAG
297 alpha diversity from unamended samples remained relatively constant throughout the
298 incubation, while in catechin-amended samples alpha diversity decreased initially, stabilized
299 through day 21, and increased by day 35 (**Fig S3A**). Beta diversity estimates revealed that
300 prokaryotic communities diverged earlier but less dramatically (**Fig. S3B**), whereas viral
301 communities exhibited a delayed but more pronounced shift in response to catechin amendment
302 (**Fig. 2B**). Specifically, for both treatments the viral response was similar up to and including day
303 7, whereas the prokaryotic community demonstrated a treatment-specific response already by
304 day 7. These different patterns suggest that viral community expression changes were not
305 strictly coupled to the prokaryotic community expression. We posit that this suggests differential
306 ecological pressures rather than co-expression of provirus within host genomes, presumably
307 with delayed viral activity stemming from relatively small day 7 prokaryote differences
308 insufficient to induce a viral community shift and lagging viral latent periods.
309 Previous perturbation studies in rewetted soils demonstrated a subset of responding
310 vOTUs dominated viral community activity, in sync with enriched taxa that were predicted to be
311 hosts [61, 63]. Given these findings, and our observations above, we hypothesized a subset of
312 newly dominant viruses may have driven the catechin-amended treatment shift in viral
313 community activity from day 7 to 14 samples, presumably by infecting previously identified
314 catechin-enriched microbes (**Fig. 1B**).
315

## 316 Viral community expression was dominated by vOTUs predicted to infect 317 key carbon cyclers

318 We predicted hosts for active vOTUs using a database augmented with sample-specific
319 MAGs and iPHoP, a machine learning based host prediction tool that aggregates many *in silico*
320 approaches into a single probabilistically modeled prediction [27]. This generated an infection
321 network (**Supplemental Data 5 & 6**) derived from 41.5% (374/900) of active vOTUs with host
322 predictions spanning 56 transcriptionally active genera and 13 phyla. With these host
323 predictions, we first explored whether there might be any signal for virus-targeted methane-
324 suppression, either through (i) increased viral activity coupled with decreased methanogen MAG
325 activity, or (ii) auxiliary metabolic gene expression. While few methanogen viruses have been
326 experimentally identified [64], recent studies have attempted to identify them with CRISPR
327 spacers and host-derived auxiliary metabolic genes [65, 66], and hinted at their potential to

328 impact methanogen metabolism [66]. In our study, active vOTUs were predicted to infect
329 *Methanothrix*, *Methanosarcina,* and *Methanobacterium_A* (**Fig. S4A**), and were generally more
330 active in unamended vs catechin-amended samples (**Fig. S4B**). Despite detecting actively-lytic
331 vOTUs predicted to infect methanogens, no clear treatment-specific infection dynamics were
332 observed. This absence of discernible patterns may reflect the relatively low abundance or slow
333 growth, manifest in sparse metatranscriptome sequencing coverage, of methanogens and their
334 viruses within the community. Additionally, we did not detect any auxiliary metabolic genes
335 within these vOTUs (**Supplementary Data 8**). Overall, any impacts methanogen viruses might
336 have had were not detectable in our analyses from these bulk metatranscriptomic data.
337 Beyond methanogens, we did observe that most viral community expression that could
338 be linked to an active MAG; in fact, this signal was dominated by vOTUs predicted to infect
339 other key carbon cyclers, with distinct responses observed in catechin-amended samples (**Fig.**
340 **2C**). In unamended samples, the bulk of viral community expression was from vOTUs predicted
341 to infect *Clostridium* and genera previously identified as catechin-sensitive [17] (i.e., prokaryotes
342 whose abundance decreased following catechin amendment) such as *Paludibacter*. In contrast,
343 viral community expression in catechin-amended samples was dominated by viruses predicted
344 to infect undescribed *Bacillota* genus JAGFXR01, *Clostridium*, and *Pseudomonas_E*, which
345 were all previously identified as key members of catechin degradation processes [17] (**Fig. 1B**).
346 Both treatments showed Clostridium vOTU dominance at day 7, but diverged by day 14.
347 JAGFXR01-infecting vOTUs surged in catechin-amended samples by day 14 and remained
348 elevated, while *Pseudomonas_E* viruses showed a transient bloom at day 21. These patterns
349 mirror those observed in the initial diversity analyses (**Fig. 2A, 2B**) with nearly all of the four-fold
350 increase in the metatranscriptomic read mappings of detected viruses from day 7 to 14 (**Fig.**
351 **2A**) attributable to predicted JAGFXR01-infecting vOTUs (**Fig. S5**).
352 Several vOTUs were highly active in specific treatments, suggesting potential ecological
353 importance, but lacked active host MAG predictions (**Fig. 2C**, no_assignment). These vOTUs
354 could have been predicted to infect inactive MAGs or iPHoP reference database MAGs, or did
355 not contain enough genetic information (as most were 5-10kb) to make a confident prediction.
356 These vOTUs might also be genetically related to other recovered vOTUs that did have host
357 predictions, and while genetic similarity does not mean they infect the same host [67, 68], it has
358 been observed for some lineages that viruses and their hosts co-evolve, and viruses that infect
359 the same host may share genes more frequently with each other than viruses that infect
360 different hosts [69, 70]. Therefore, host predictions for vOTUs genetically related to the highly
361 active vOTUs without host predictions may give an idea about their potential host target. Based
362 on this, we explored the potential hosts of these highly active vOTUs through (i) predictions to
363 inactive or reference MAGs in the iPHoP host database, and (ii) taxonomic similarity through
364 gene-sharing networks (see Methods) to other vOTUs with a host prediction. Towards the first
365 approach, the two most abundant unassigned vOTUs in unamended samples were predicted to
366 infect *Lachnoclostridium* reference MAGs and another previously unassigned vOTU was
367 predicted to infect *Pseudomonas_E* reference MAGs (Supplementary Data 6**, Fig. S6**). Towards
368 the second approach, all but one unassigned vOTU clustered with at least one other vOTU
369 recovered in this study with a host prediction. Within each vOTU infection network, all vOTUs
370 with a host prediction were predicted to infect the same host genus (**Fig. S6**), of which
371 *Pseudomonas_E*, JAGFXR01, and JAEXAI01 were previously identified as polyphenol

degrading, catechin-enriched taxa [17]. Together, these relationships suggest these highly active vOTUs infected ecologically relevant hosts and that viral predation of key carbon cyclers may be even higher than predicted above. While viral taxonomy is not an explicit feature used in virus-host prediction algorithms, future methods could benchmark the usefulness of viral taxonomy in virus-host predictions at different taxonomic ranks.

## Differential virus responses within key carbon cycler infection networks suggests vOTU-specific niches

A subset of MAGs from each genus of key carbon cyclers responded to catechin amendment [17], implying approximately species-level differential responses that might correspond to distinct niches, as observed in other systems [71, 72]. Thus, we posited vOTUs predicted to infect these key carbon cyclers would also differentially respond, indicative of distinct vOTU-level niches [73, 74]. To test this, we employed indicator species analysis [32] as a statistical framework to determine which vOTUs differentially responded across sample clusters formed by affinity propagation, which roughly clustered by treatment and timepoint (**Fig. 2B**). Indicator species analysis revealed 121 indicator vOTUs for at least one combination of sample clusters (*P* value ≤ 0.01, **Fig. 2D**). Within the key carbon cycler infection networks, we observed indicator vOTUs for five (for *Clostridium*), two (for *Pseudomonas_E*), two (for JAGFXR01), and one (for *Paludibacter*) different sample cluster combinations, demonstrating differential responses between treatments and/or over time within infection networks (**Fig. 2D**).

Given such varied sub-genus-level vOTU and MAG responses, we explored whether ecological patterning could refine virus-host predictions below the predicted genus level. Virus-host linkages at genus and species ranks are known to be challenging as genome-based prediction methods (e.g., iPHoP [27]) differed greatly compared to experimentally benchmarked approaches such as proximity ligation [75]. Correlation-based inference is also challenging in complex communities, particularly on time-series data [76]. However, we reasoned that coupling genome-based host predictions and correlated co-occurrence dynamics might provide independent evidence to distinguish active infections among multiple potential hosts at the species level — at least for the subset of taxa where time-resolved patterning was distinct across time points and/or treatments (**Fig. 3A**). Using this rationale, we then tested whether treatment-enriched indicator vOTUs showed correlated dynamics with co-enriched predicted hosts, focusing on the four key carbon cycler infection networks identified above.

In three of the four key carbon cycler infection networks (JAGFXR01, *Pseudomonas_E*, and *Paludibacter*), nearly all indicator vOTUs were significantly correlated with their predicted hosts based on relative expression across all samples (**Fig. 3B, upper panel**; adjusted p-value < 0.05), qualitatively supporting the MAG-level host predictions within each genus. However, of the four genera tested, only *Pseudomonas_E*-infecting indicator vOTUs were significantly more correlated with predicted MAGs versus non-predicted MAGs in the same host genus, likely because many MAGs within each genus responded similarly across samples (**Fig. S7, dashed lines**). *Clostridium*-infecting vOTUs were only weakly correlated to their predicted host MAGs, and correlation distributions were nearly identical between predicted and non-predicted *Clostridium* MAGs and predicted vOTUs (**Fig. 3B**). Within each set of vOTU-MAG predictions for the four key carbon cyclers minus *Clostridium*, indicator vOTUs were significantly correlated with at least one predicted host MAG (**Fig. S8**). These observations suggest that among

multiple MAGs within a genus predicted as potential hosts for a treatment-enriched vOTU, co-enriched MAGs likely represent the primary targets sustaining viral population activity, though infection of other predicted hosts cannot be excluded. While previous paired virus-prokaryote eco-genomics studies in soil ecosystems, including at this site, have documented correlations between vOTU abundance and their hosts across samples [77–79], our approach extends this framework by seeking to resolve sub-genus vOTU-MAG relationships. If future experimental measurements can be developed to confirm these relationships, such analytics would allow for community-wide increasing precision for inferring active infection networks within complex systems.

## One dominant vOTU may facilitate rewiring of community carbon metabolism

One vOTU (contig_591846, 40 kb) was predicted to infect JAGFXR01 and dominated viral community expression in catechin day 14-35 samples. Quantitatively, this vOTU alone accounted for 96-100% of total detected JAGFXR01-infecting vOTU activity (**Fig. 2C**) and its expression was on average 82-308x higher than that of the next most highly expressed vOTU predicted to infect JAGFXR01 (**Fig. 4A**, upper panel). Genomically, contig_591846 was (i) high-quality (>90% estimated completeness), but not complete (no direct terminal repeats or circularization), (ii) appeared to be a double-stranded DNA, encapsidated, tailed, temperate phage (**Fig. 4B,** lower panel), and (iii) shared less than 10% of its proteins with other JAGFXR01-infecting vOTUs (**Fig. S9**). Taxonomically, gene-sharing analyses against vConTACT3 [38] virus reference database v228 showed that contig_591846 clustered with eight viral sequences at the family-level (an unnamed family within phylum *Caudoviracetes*), six of which were RefSeq viruses belonging to genera *Abouovirus* and *Jimmervirus*, and two vOTUs identified in this study (**Supplementary Data 5**). Together these findings suggest that the dominantly responding contig_591846 was near-complete, genetically divergent from previously identified viruses, and likely an induced prophage lytically responding to the catechin-enriched JAGFXR01 niche.

To determine if contig_591846 was undergoing induced lysis, we assessed its metatranscriptome expression of functionally relevant lysis and lysogeny genes from days 14, 21, and 35 catechin-amended samples, and its genomic copy number compared to that of its predicted JAGFXR01 host MAG. The former revealed that contig_591846 in catechin-amended samples (i) consistently expressed structural genes, (ii) structural gene expression was two- to three-fold higher than the integrase region, and (iii) antirepressor gene expression was two- to three-fold higher than that of the cI repressor (**Fig. 4B**). These observations are consistent with signature gene expression expectations for lytically-active phages [80]. The latter revealed genomic representation (where metagenomes were available) that ranged from 20- to 156-fold more abundant than its predicted JAGFXR01 host MAG (**Fig. 4A**, lower panel).

Metagenome-based virus-host microbe ratios vary widely [81], and are not typically reported at the vOTU-MAG level due to prediction tool imprecision (discussed above). Thus, we treated this unique treatment-responder situation as an opportunity to push biological inferences, with a focus on attempting to genomically assess the average per-cell viral burst size for this lineage. Specifically, JAGFXR01 vOTU-MAG relative abundance ratios >100 from days 14-35 suggest a bust size of at least this, which we posit to indicate that JAGFXR01 is

459 undergoing intense, sustained viral replication.
460 Given this evidence for intensive lytic infection pressure, we wondered what downstream
461 effects JAGFXR01 lysis might have on the broader microbial community. Previous soil virus
462 studies experimentally demonstrated that viruses alleviate some microbial nutrient limitations
463 [28], perhaps through frequently lysing abundant hosts to increase diversity via auxiliary
464 community members scavenging nutrient-rich lysates. In our system, we reasoned that
465 enrichment of catechin-degrading JAGFXR01 cells would open a new niche for opportunistic
466 viruses, generating a “kill-the-winner” [82] scenario for responding viruses, where viral predation
467 rapidly increases for abundant hosts as a top-down control, and this in turn could stimulate
468 community diversity via lysis-released partially metabolized catechin intermediates diversifying
469 resources available to non-catechin-degrading microbes. Indeed, we previously showed
470 aromatic catechin degradation intermediates accumulated in catechin-amended sample pore
471 waters [17]. However, this left open the question of their bioavailability and what taxa might be
472 utilizing them.
473 To investigate whether non-catechin-degrading MAGs were utilizing these downstream
474 intermediates, we assessed the relative expression of genes encoding catechin-degradation
475 intermediate metabolism, specifically phloroglucinol transformation (*PGR*, *pgthAB*) in MAGs that
476 were not detected to directly degrade catechin. Of the 104 MAGs expressing genes for aromatic
477 intermediate metabolism but not catechin degradation, 25 MAGs (21 genera) showed elevated
478 expression of at least one of these phloroglucinol transformation genes in catechin-amended
479 samples at one or more timepoints during the virus-active period (days 14-35), with 7 MAGs (5
480 genera) elevated at two timepoints and 2 MAGs (2 genera) elevated at all timepoints (**Fig. 5A**).
481 Furthermore, hydrogenase activity among these MAGs was elevated in catechin-amended
482 samples, suggesting hydrogen consumption was maintained (**Fig. 5B**). We see two possible
483 explanations for these observations. Our favored hypothesis is that phage-mediated lysis
484 released aromatic intermediates that became bioavailable in the environment to stimulate
485 intermediate-specific metabolic activity, and presumably growth, by non-catechin degraders
486 (**Fig. 5C**). However, it is also possible that additional polyphenol degradation pathways, not
487 detected in our high-quality MAGs, were actually present in the non-catechin-degraders and
488 became upregulated in catechin-amended samples due to altered resource availability or redox
489 state. Though stable-isotope probing experiments are incredibly rare in complex communities,
490 especially for viruses, due to numerous challenges [64, 83, 84], future microcosm experiments
491 amended with $^{13}$C-catechin could evaluate these omics-enabled hypotheses about which taxa
492 utilize catechin-degradation intermediates, and which viruses infect them.
493

494 ## Conclusion

495 Overall, these data support a model where intense viral predation of the dominant catechin
496 degrader acts as a viral shunt that redirects partially metabolized substrates to the broader
497 microbial community. In this system, the net effect is that hydrogen consumption is maintained
498 (via downstream phloroglucinol degradation) such that monopolization of catechin-derived
499 carbon by a single taxon is prevented. A single vOTU, predicted to infect a catechin-degrading
500 microbe, became highly enriched alongside its host and genomic copy numbers >100-fold more
501 abundant than that of its predicted host, were interpreted to represent high lineage-specific lysis
502 rates. As such, these findings extend the microbe-focused view of metabolic rewiring described

previously [17] by demonstrating that carbon rewiring during catechin amendment is shaped not only by microbial competition and lineage-specific metabolic thermodynamic constraints, but also by a second layer of biological control imposed by viruses. The addition of a viral layer to the original model reveals how viral shunting can redistribute catechin-derived carbon to downstream taxa that presumably outcompete methanogens for key substrates to depress methanogenesis. This viral dimension is not simply an overlay on microbial processes; it introduces new causal pathways—through predation, population regulation, and the release of partially transformed metabolites—that redistribute resources and shuffle the deck on which taxa benefit from catechin amendment. This integrated virus–microbe framework thus offers a more complete mechanistic understanding of how rapid substrate perturbations drive community-level metabolic transitions. As calls grow to document climate-driven changes in microbial diversity [85], our findings point to a broader principle with relevance beyond these peatland microcosms: viral regulation of microbial metabolism is a fundamental driver of carbon and nutrient flow across ecosystems, shaping how microbial communities reorganize under environmental change.

## Acknowledgements

We are grateful to the EMERGE Biology Integration Institute's field team for collecting peat samples that were used in the bioreactors, as well as years of dedicated field support to make the longer-term ecosystem-scale data available. We also thank members of the Sullivan Lab for constructive conversations and suggestions through several years of lab meeting presentations, John Bouranis for early project meeting suggestions, and Simon Roux for troubleshooting iPHoP custom database construction.

## Author contributions

**James Riddell V** (Conceptualization, Data curation, Formal analysis, Investigation, Funding acquisition, Methodology, Project administration, Software, Validation, Visualization, Writing- original draft preparation, Writing- review and editing), **Rokaiya Nurani Shatadru** (Conceptualization, Validation, Methodology, Visualization, Writing — drafting short sections, Writing — review & editing), **Garrett J Smith** (Conceptualization, Validation, Methodology, Writing — review & editing), **Bridget B McGivern** (Conceptualization, Data curation, Methodology, Writing — review & editing), **Jared B Ellenbogen** (Conceptualization, Investigation, Writing — review & editing), **Sophie K Jurgensen** (Conceptualization, Writing — review & editing), **Ami Fofana** (Visualization, Writing — review & editing), **Malak M Tfaily** (Conceptualization, Funding Acquisition, Project administration, Supervision, Writing — review & editing), **Kelly C Wrighton** (Conceptualization, Funding Acquisition, Project administration, Supervision, Writing — drafting short sections, Writing — review & editing), **Matthew B Sullivan** (Conceptualization, Funding Acquisition, Methodology, Project administration, Resources, Supervision, Writing — drafting short sections, Writing — review & editing).

## Supplementary data

Supplementary Figures
Supplementary Data

## Conflicts of interest

The authors declare no competing financial interests.

## Funding

Funding supporting this work includes a National Science Foundation (NSF) Graduate Research Fellowship Program (Grant No. DGE-1343012) to JRV, a Department of Energy Biological and Environmental Research Award (#DE-SC0023307) to MBS and MMT, a Grantham Foundation for the Protection of the Environment Award to MBS, KCW, MMT, and an NSF Biology Integration Institutes Program Award (no. 2022070) to the EMERGE Biology Integration Institute.

The work (proposal: https://doi.org/10.46936/10.25585/60001190) conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. Analyses were performed using computing resources partially subsidized by the Ohio Supercomputer Center.

## Data availability

Metatranscriptomic and metagenomic reads used in this current study are available at NCBI under BioProjectID PRJNA1137330. See Supplementary Data 1 for specific accession numbers. All code required to reproduce the analyses and visualizations used in this publication are available at https://github.com/jamesriddellv/Grantham_Bioreactor, and are described in the README file. Supplementary Data 1 through 9 are available in Supplementary Data. Assembled contigs and viral contigs identified in this study are available at https://zenodo.org/records/17652154 [86]. Sample metadata and intermediate coverage files used in this study that were not included in the supplementary data are available at: https://zenodo.org/records/17652194 [87].

## References

1. Cavicchioli R et al. Scientists' warning to humanity: microorganisms and climate change. *Nat Rev Microbiol* 2019;**17**:569–586. https://doi.org/10.1038/s41579-019-0222-5
2. Naga NG et al. The silent microbial shift: climate change amplifies pathogen evolution, microbiome dysbiosis, and antimicrobial resistance. *Trop Dis Travel Med Vaccines* 2025;**11**:43. https://doi.org/10.1186/s40794-025-00275-y
3. Baker DM et al. Climate change promotes parasitism in a coral symbiosis. *ISME J* 2018;**12**:921–930. https://doi.org/10.1038/s41396-018-0046-8
4. Rudgers JA et al. Climate Disruption of Plant-Microbe Interactions. *Annu Rev Ecol Evol Syst* 2020;**51**:561–586. https://doi.org/10.1146/annurev-ecolsys-011720-090819
5. Addison SL et al. Partner or perish: tree microbiomes and climate change. *Trends Plant Sci* 2024;**29**:1029–1040. https://doi.org/10.1016/j.tplants.2024.03.008

6. Jansson JK, Hofmockel KS. Soil microbiomes and climate change. *Nat Rev Microbiol* 2020;**18**:35–46. https://doi.org/10.1038/s41579-019-0265-7

7. Feng J et al. Warming-induced permafrost thaw exacerbates tundra soil carbon decomposition mediated by microbial community. *Microbiome* 2020;**8**:3. https://doi.org/10.1186/s40168-019-0778-3

8. Xue K et al. Tundra soil carbon is vulnerable to rapid microbial decomposition under climate warming. *Nat Clim Change* 2016;**6**:595–600. https://doi.org/10.1038/nclimate2940

9. Ramage J et al. The Net GHG Balance and Budget of the Permafrost Region (2000–2020) From Ecosystem Flux Upscaling. *Glob Biogeochem Cycles* 2024;**38**:e2023GB007953. https://doi.org/10.1029/2023GB007953

10. Varner RK et al. Permafrost thaw driven changes in hydrology and vegetation cover increase trace gas emissions and climate forcing in Stordalen Mire from 1970 to 2014. *Philos Trans R Soc Math Phys Eng Sci* 2021;**380**:20210022. https://doi.org/10.1098/rsta.2021.0022

11. Holmes ME et al. Carbon Accumulation, Flux, and Fate in Stordalen Mire, a Permafrost Peatland in Transition. *Glob Biogeochem Cycles* 2022;**36**:e2021GB007113. https://doi.org/10.1029/2021GB007113

12. Woodcroft BJ et al. Genome-centric view of carbon processing in thawing permafrost. *Nature* 2018;**560**:49–54. https://doi.org/10.1038/s41586-018-0338-1

13. Jansson JK, McClure R, Egbert RG. Soil microbiome engineering for sustainability in a changing environment. *Nat Biotechnol* 2023;**41**:1716–1728. https://doi.org/10.1038/s41587-023-01932-3

14. Peixoto R et al. Microbial solutions must be deployed against climate catastrophe. *Npj Biofilms Microbiomes* 2024;**10**:122. https://doi.org/10.1038/s41522-024-00591-9

15. Olsrud M, Michelsen A. Effects of shading on photosynthesis, plant organic nitrogen uptake, and root fungal colonization in a subarctic mire ecosystem. *Botany* 2009;**87**:463–474. https://doi.org/10.1139/B09-021

16. Ben Y et al. Metabolomic indicators anticipate climate-driven permafrost ecosystem collapse before visible change. 2025. Rochester, NY: Social Science Research Network, 2025.

17. McGivern BB et al. Polyphenol rewiring of the microbiome reduces methane emissions. *ISME J* 2025;wraf108. https://doi.org/10.1093/ismejo/wraf108

18. Jover LF et al. The elemental composition of virus particles: implications for marine biogeochemical cycles. *Nat Rev Microbiol* 2014;**12**:519–528. https://doi.org/10.1038/nrmicro3289

19. Howard-Varona C et al. Environment-specific virocell metabolic reprogramming. *ISME J* 2024;**18**:wrae055. https://doi.org/10.1093/ismejo/wrae055

20. Dalcin Martins P et al. Viral and metabolic controls on high rates of microbial sulfur and carbon cycling in wetland ecosystems. *Microbiome* 2018;**6**:138. https://doi.org/10.1186/s40168-018-0522-4

21. Ji M et al. Tundra Soil Viruses Mediate Responses of Microbial Communities to Climate Warming. *mBio* 2023;**0**:e03009-22. https://doi.org/10.1128/mbio.03009-22

22. Emerson JB et al. Host-linked soil viral ecology along a permafrost thaw gradient. *Nat Microbiol* 2018;**3**:870–880. https://doi.org/10.1038/s41564-018-0190-y
23. Zimmerman AE et al. Estimating the Importance of Viral Contributions to Soil Carbon Dynamics. *Glob Change Biol* 2024;**30**:e17524. https://doi.org/10.1111/gcb.17524
24. Albright MBN et al. Experimental evidence for the impact of soil viruses on carbon cycling during surface plant litter decomposition. *ISME Commun* 2022;**2**:1–8. https://doi.org/10.1038/s43705-022-00109-4
25. Talmy D et al. Viruses in multi-scale ocean models: challenges and opportunities. *Front Mar Sci* 2025;**12**. https://doi.org/10.3389/fmars.2025.1717845
26. Camargo AP et al. Identification of mobile genetic elements with geNomad. *Nat Biotechnol* 2023;1–10. https://doi.org/10.1038/s41587-023-01953-y
27. Roux S et al. iPHoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. *PLOS Biol* 2023;**21**:e3002083. https://doi.org/10.1371/journal.pbio.3002083
28. Tong D et al. Viral lysing can alleviate microbial nutrient limitations and accumulate recalcitrant dissolved organic matter components in soil. *ISME J* 2023;**17**:1247–1256. https://doi.org/10.1038/s41396-023-01438-5
29. Chaumeil P-A et al. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 2022;**38**:5315–5316. https://doi.org/10.1093/bioinformatics/btac672
30. Shaffer M et al. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res* 2020;**48**:8883–8900. https://doi.org/10.1093/nar/gkaa621
31. McGivern BB et al. Microbial polyphenol metabolism is part of the thawing permafrost carbon cycle. *Nat Microbiol* 2024;**9**:1454–1466. https://doi.org/10.1038/s41564-024-01691-0
32. Smid M et al. Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. *BMC Bioinformatics* 2018;**19**:236. https://doi.org/10.1186/s12859-018-2246-7
33. Sun CL et al. Virus ecology and 7-year temporal dynamics across a permafrost thaw gradient. *Environ Microbiol* 2024;**26**:e16665. https://doi.org/10.1111/1462-2920.16665
34. Nayfach S et al. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat Biotechnol* 2021;**39**:578–585. https://doi.org/10.1038/s41587-020-00774-7
35. Roux S et al. Minimum Information about an Uncultivated Virus Genome (MIUViG). *Nat Biotechnol* 2019;**37**:29–37. https://doi.org/10.1038/nbt.4306
36. Hyatt D et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;**11**:119. https://doi.org/10.1186/1471-2105-11-119
37. Coclet C, Camargo AP, Roux S. MVP: a modular viromics pipeline to identify, filter, cluster, annotate, and bin viruses from metagenomes. *mSystems* 2024;**9**:e00888-24. https://doi.org/10.1128/msystems.00888-24

38. Bolduc B. vConTACT3: Viral Genome Clustering and Taxonomy Tool. 2025. 2025.
39. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–359. https://doi.org/10.1038/nmeth.1923
40. Danecek P et al. Twelve years of SAMtools and BCFtools. *GigaScience* 2021;**10**:giab008. https://doi.org/10.1093/gigascience/giab008
41. Bushnell B. BBMap: A Fast, Accurate, Splice-Aware Aligner. 2014.
42. Gregory AC et al. Marine DNA Viral Macro- and Microdiversity from Pole to Pole. *Cell* 2019;**177**:1109-1123.e14. https://doi.org/10.1016/j.cell.2019.03.040
43. Muscatt G et al. Crop management shapes the diversity and activity of DNA and RNA viruses in the rhizosphere. *Microbiome* 2022;**10**:181. https://doi.org/10.1186/s40168-022-01371-3
44. Coclet C et al. Virus diversity and activity is driven by snowmelt and host dynamics in a high-altitude watershed soil ecosystem. *Microbiome* 2023;**11**:237. https://doi.org/10.1186/s40168-023-01666-z
45. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma Oxf Engl* 2010;**26**:841–842. https://doi.org/10.1093/bioinformatics/btq033
46. Rideout JR et al. scikit-bio/scikit-bio: scikit-bio 0.7.1.post1. 2025. Zenodo, 2025.
47. Oksanen J et al. vegan: Community Ecology Package. 2025.
48. Bodenhofer U, Kothmeier A, Hochreiter S. APCluster: an R package for affinity propagation clustering. *Bioinformatics* 2011;**27**:2463–2464. https://doi.org/10.1093/bioinformatics/btr406
49. Shannon P et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* 2003;**13**:2498–2504. https://doi.org/10.1101/gr.1239303
50. Pratama AA et al. Expanding standards in viromics: in silico evaluation of dsDNA viral genome identification, classification, and auxiliary metabolic gene curation. *PeerJ* 2021;**9**:e11447. https://doi.org/10.7717/peerj.11447
51. Tian F et al. Prokaryotic-virus-encoded auxiliary metabolic genes throughout the global oceans. *Microbiome* 2024;**12**:159. https://doi.org/10.1186/s40168-024-01876-z
52. De Cáceres M, Legendre P, Moretti M. Improving indicator species analysis by combining groups of sites. *Oikos* 2010;**119**:1674–1684. https://doi.org/10.1111/j.1600-0706.2010.18334.x
53. Aroney STN et al. CoverM: read alignment statistics for metagenomics. *Bioinformatics* 2025;**41**:btaf147. https://doi.org/10.1093/bioinformatics/btaf147
54. Moraru C. VirClust—A Tool for Hierarchical Clustering, Core Protein Detection and Annotation of (Prokaryotic) Viruses. *Viruses* 2023;**15**:1007. https://doi.org/10.3390/v15041007
55. Noguchi H, Taniguchi T, Itoh T. MetaGeneAnnotator: Detecting Species-Specific Patterns of Ribosomal Binding Site for Precise Gene Prediction in Anonymous Prokaryotic and Phage Genomes. *DNA Res* 2008;**15**:387–396. https://doi.org/10.1093/dnares/dsn027
56. Hackl T et al. gggenomes: effective and versatile visualizations for comparative genomics. 2024. arXiv, 2024.

718 57. Wilke CO. ggridges: Ridgeline Plots in ‘ggplot2’. 2017. 2017. , 0.5.7
719 58. Roux S, Coclet C. Viromics approaches for the study of viral diversity and
720 ecology in microbiomes. *Nat Rev Genet* 2025;1–15. https://doi.org/10.1038/s41576-025-
721 00871-w
722 59. Marcinowski L et al. Real-time Transcriptional Profiling of Cellular and Viral Gene
723 Expression during Lytic Cytomegalovirus Infection. *PLOS Pathog* 2012;**8**:e1002908.
724 https://doi.org/10.1371/journal.ppat.1002908
725 60. Roux S et al. Towards quantitative viromics for both double-stranded and single-
726 stranded DNA viruses. *PeerJ* 2016;**4**:e2777. https://doi.org/10.7717/peerj.2777
727 61. Santos-Medellín C et al. Viral but not bacterial community successional patterns
728 reflect extreme turnover shortly after rewetting dry soils. *Nat Ecol Evol* 2023;**7**:1809–1822.
729 https://doi.org/10.1038/s41559-023-02207-5
730 62. Santos-Medellín C et al. Spatial turnover of soil viral populations and genotypes
731 overlain by cohesive responses to moisture in grasslands. *Proc Natl Acad Sci*
732 2022;**119**:e2209132119. https://doi.org/10.1073/pnas.2209132119
733 63. Nicolas AM et al. A subset of viruses thrives following microbial resuscitation
734 during rewetting of a seasonally dry California grassland soil. *Nat Commun* 2023;**14**:5835.
735 https://doi.org/10.1038/s41467-023-40835-4
736 64. Ngo VQH et al. Diversity of novel archaeal viruses infecting methanogens
737 discovered through coupling of stable isotope probing and metagenomics. *Environ*
738 *Microbiol* 2022;**24**:4853–4868. https://doi.org/10.1111/1462-2920.16120
739 65. Medvedeva S et al. A compendium of viruses from methanogenic archaea
740 reveals their diversity and adaptations to the gut environment. *Nat Microbiol* 2023;**8**:2170–
741 2182. https://doi.org/10.1038/s41564-023-01485-w
742 66. Wang L et al. Potential metabolic and genetic interaction among viruses,
743 methanogen and methanotrophic archaea, and their syntrophic partners. *ISME Commun*
744 2022;**2**:1–10. https://doi.org/10.1038/s43705-022-00135-2
745 67. Sullivan MB et al. Genomic analysis of oceanic cyanobacterial myoviruses
746 compared with T4-like myoviruses from diverse hosts and environments. *Environ Microbiol*
747 2010;**12**:3035–3056. https://doi.org/10.1111/j.1462-2920.2010.02280.x
748 68. Geoghegan JL, Duchêne S, Holmes EC. Comparative analysis estimates the
749 relative frequencies of co-divergence and cross-species transmission within viral families.
750 *PLOS Pathog* 2017;**13**:e1006215. https://doi.org/10.1371/journal.ppat.1006215
751 69. Bellec L et al. Cophylogenetic interactions between marine viruses and
752 eukaryotic picophytoplankton. *BMC Evol Biol* 2014;**14**:59. https://doi.org/10.1186/1471-
753 2148-14-59
754 70. Ignacio-Espinoza JC, Sullivan MB. Phylogenomics of T4 cyanophages: lateral
755 gene transfer in the ‘core’ and origins of host genes. *Environ Microbiol* 2012;**14**:2113–
756 2126. https://doi.org/10.1111/j.1462-2920.2012.02704.x
757 71. Orellana LH et al. Niche differentiation among annually recurrent coastal Marine
758 Group II Euryarchaeota. *ISME J* 2019;**13**:3024–3036. https://doi.org/10.1038/s41396-019-
759 0491-z

72. Reji L, Zhang X. Genome-Resolved Metagenomics Informs the Functional Ecology of Uncultured Acidobacteria in Redox Oscillated Sphagnum Peat. *mSystems* 2022. https://doi.org/10.1128/msystems.00055-22

73. Duhaime MB et al. Comparative Omics and Trait Analyses of Marine Pseudoalteromonas Phages Advance the Phage OTU Concept. *Front Microbiol* 2017;**8**. https://doi.org/10.3389/fmicb.2017.01241

74. Gregory AC et al. Genomic differentiation among wild cyanophages despite widespread horizontal gene transfer. *BMC Genomics* 2016;**17**:930. https://doi.org/10.1186/s12864-016-3286-x

75. Shatadru RN et al. Benchmarking with synthetic communities provides a baseline for virus-host inferences from Hi-C proximity linking. *PLOS Biol* 2025;**23**:e3003510. https://doi.org/10.1371/journal.pbio.3003510

76. Coenen AR, Weitz JS. Limitations of Correlation-Based Inference in Complex Virus-Microbe Communities. *mSystems* 2018. https://doi.org/10.1128/msystems.00084-18

77. Liao H et al. Mesophilic and thermophilic viruses are associated with nutrient cycling during hyperthermophilic composting. *ISME J* 2023;**17**:916–930. https://doi.org/10.1038/s41396-023-01404-1

78. Du S et al. Highly host-linked viromes in the built environment possess habitat-dependent diversity and functions for potential virus-host coevolution. *Nat Commun* 2023;**14**:2676. https://doi.org/10.1038/s41467-023-38400-0

79. Gios E et al. High niche specificity and host genetic diversity of groundwater viruses. *ISME J* 2024;**18**:wrae035. https://doi.org/10.1093/ismejo/wrae035

80. Yang H et al. Transcription regulation mechanisms of bacteriophages. *Bioengineered* 2014;**5**:300–304. https://doi.org/10.4161/bioe.32110

81. López-García P et al. Metagenome-derived virus-microbe ratios across ecosystems. *ISME J* 2023;**17**:1552–1563. https://doi.org/10.1038/s41396-023-01431-y

82. Thingstad TF. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanogr* 2000;**45**:1320–1328. https://doi.org/10.4319/lo.2000.45.6.1320

83. Trubl G et al. Active virus-host interactions at sub-freezing temperatures in Arctic peat soil. *Microbiome* 2021;**9**:208. https://doi.org/10.1186/s40168-021-01154-2

84. Greenlon A et al. Quantitative Stable-Isotope Probing (qSIP) with Metagenomics Links Microbial Physiology and Activity to Soil Moisture in Mediterranean-Climate Grassland Ecosystems. *mSystems* 2022;**7**:e00417-22. https://doi.org/10.1128/msystems.00417-22

85. Gilbert JA et al. Safeguarding microbial biodiversity: microbial conservation specialist group within the species survival commission of the International Union for Conservation of Nature. *FEMS Microbiol Ecol* 2025;**101**:fiaf107. https://doi.org/10.1093/femsec/fiaf107

86. Riddell V J. Assemblies and vOTUs from Riddell V et al. zenodo.

87. Riddell V J. Data repository for modules 02 and 03 in Riddell V et al. zenodo.

# Figure Legends

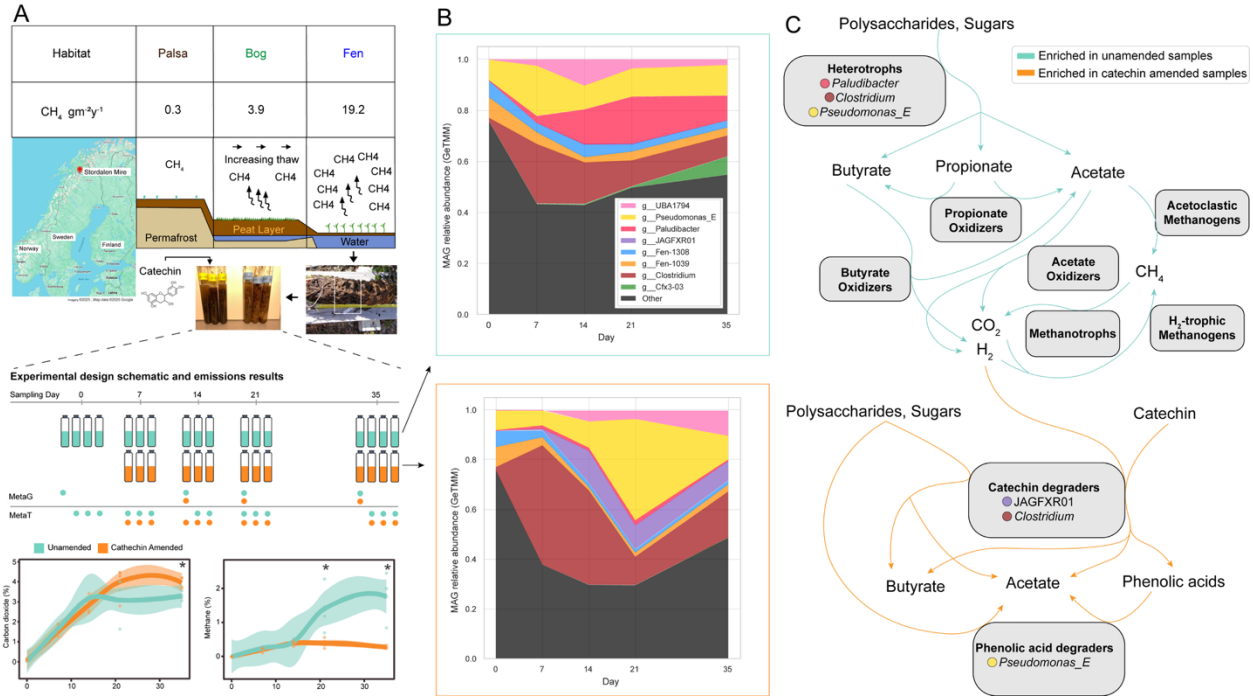



**Figure 1. Experimental design overview and summary of prokaryotic community dynamics in response to catechin amendment. (A)** Peat cores were taken from a fen in Stordalen Mire, Sweden, in July 2016. Microcosms were constructed by placing ~6g of soil from 9-19cm below the surface (black/white box) into glass Balch tubes. Catechin was added to half of the samples to a final concentration of 4.4 mM, and microcosms were destructively sampled at five timepoints. The sequencing done for each microcosm is specified in the experimental design schematic. Methane flux values derived from [11]. Experimental design schematic and gas measurements adapted [17] under the Creative Commons Attribution license (CC BY). Peat core photo taken by Nicole Raab as part of the EMERGE field team. **(B)** MAG community relative activity was visualized by taking the average GeTMM value of each MAG in each set of sample replicates and summing for all MAGs in the same genus. If a genus-level assignment was not available, the next highest resolution taxonomic rank was used to ensure all MAGs were accounted for. Taxa accounting for >5% of the total MAG relative activity in at least one set of sample replicates were plotted on stacked area charts showing average GeTMM relative activity over time separated by treatment. All other taxa were summed into “Other”. **(C)** Conceptual diagram of major prokaryotic community metabolisms in the microcosms as described previously [17]. Turquoise arrows specify metabolisms enriched in unamended samples, and orange arrows specify metabolisms enriched in catechin-amended samples.

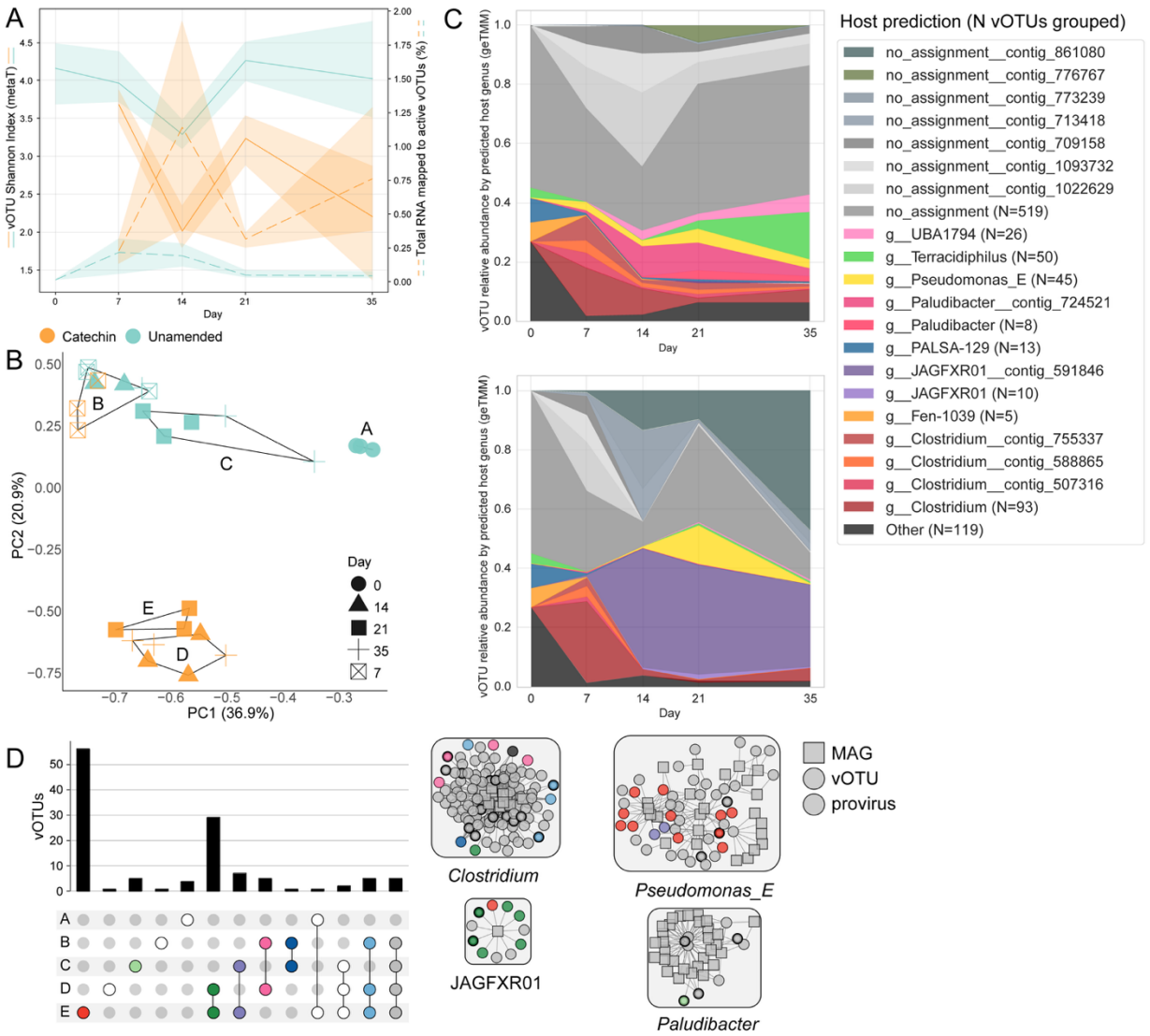



**Figure 2. vOTU communities differed between treatments and over time. (A)** vOTU Shannon diversity and proportion of reads mapped to active vOTUs. Shaded areas represent 95% confidence intervals. **(B)** Principal component analysis of Bray–Curtis dissimilarity among metatranscriptomic samples using active vOTU GeTMM relative abundance values. Samples labeled by day; colors correspond to treatment. Sample clusters determined by affinity propagation are outlined and labeled with black letters. **(C)** Stacked area chart of active vOTU GeTMM relative activity in unamended samples (top) and catechin-amended samples (bottom) aggregated by predicted host genus (sum). Aggregated genera with <5% contribution in at least one set of sample replicates were grouped into "Other". vOTUs accounting for >4% of average GeTMM relative activity in at least one set of sample replicates were plotted individually in a similar shade to their predicted host genus. **(D)** (left) Upset plot of indicator species analysis of active vOTUs across sample clusters (p-value ≤ 0.01). Sample clusters correspond to those defined in (B). (right) Infection network of key carbon cyclers. Squares represent MAGs, circles represent vOTUs, and bold circles represent proviruses. Edges represent iPHoP phage-host predictions with confidence score >90. Colored circles represent indicator vOTUs based on indicator species analysis. The circle color denotes the combination of sample clusters the

vOTU was an indicator for. Cluster combinations for which there were no indicator vOTUs in the four infections shown are marked in white for visual brevity.

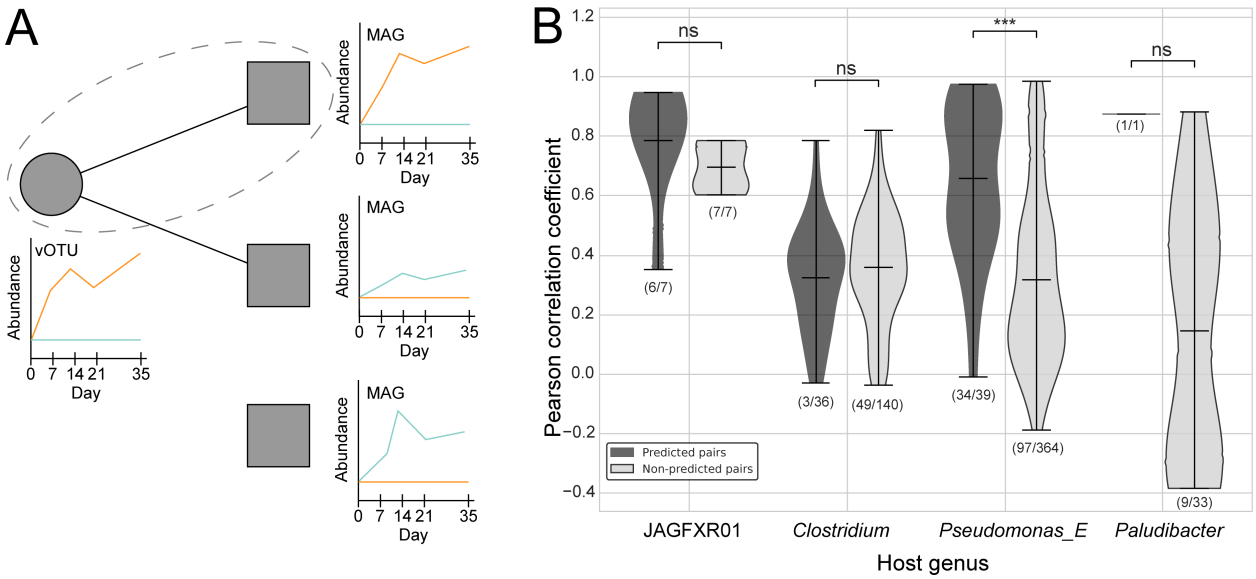



**Figure 3. Correlation inference of indicator vOTUs and predicted hosts within key carbon cycler infection networks. (A)** Conceptual overview of correlation-based inference from metatranscriptomic data. The top virus-host pair demonstrates a virus-host prediction (black line) where vOTU (circle) and MAG (square) relative abundance are correlated across samples. The middle virus-host pair represents an in silico prediction, but where the vOTU and MAG are not correlated. The last MAG is in the same genus as the other two, but the vOTU is not predicted to infect it. **(B)** Violin plot of pearson correlation coefficients between indicator vOTUs and their predicted hosts and non-predicted hosts within each genus-level infection network (from Fig. 2D). ns = not significant, *** = *P* value < 0.005. Dark grey plots represent values between iPHoP-predicted vOTU-MAG pairs, and light grey plots represent values between vOTUs and MAGs they were not predicted to infect, but were in the same genus as their predicted host MAG. Values at the bottom of each violin plot represent how many vOTU-MAG pairs were significantly correlated (BH-adjusted *P* value < 0.05) between either predicted MAGs or non-predicted MAGs in the infection network out of the total pairs making up the distribution.

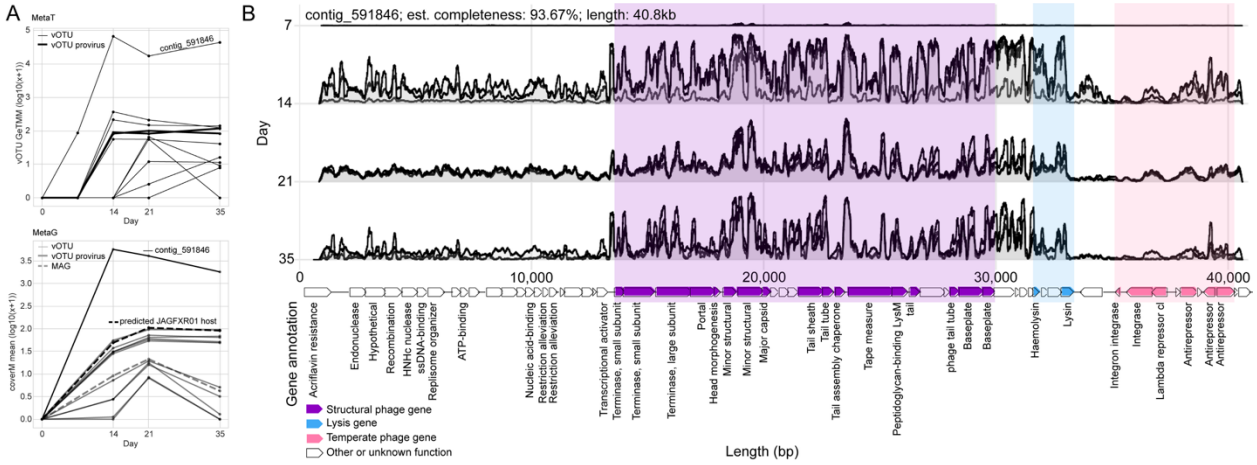



**Figure 4. Relative abundance of JAGFXR01 MAGs and predicted viruses. (A)** GeTMM values of active vOTUs predicted to infect JAGFXR01 averaged across replicate catechin-amended samples (top). Lines are colored by indicator value as in Fig. 2D. CoverM trimmed mean coverage of active JAGFXR01-infecting vOTUs and JAGFXR01 MAGs (bottom). **(B)** Contig_591846 read pileup from catechin-amended metatranscriptomic sample reads mapping >90% ANI to the coding strand. Replicate samples are plotted on top of each other. Predicted ORFs are plotted below, with abbreviated functional annotations based on DRAMv and GeNomad. Select genes are colored by function.

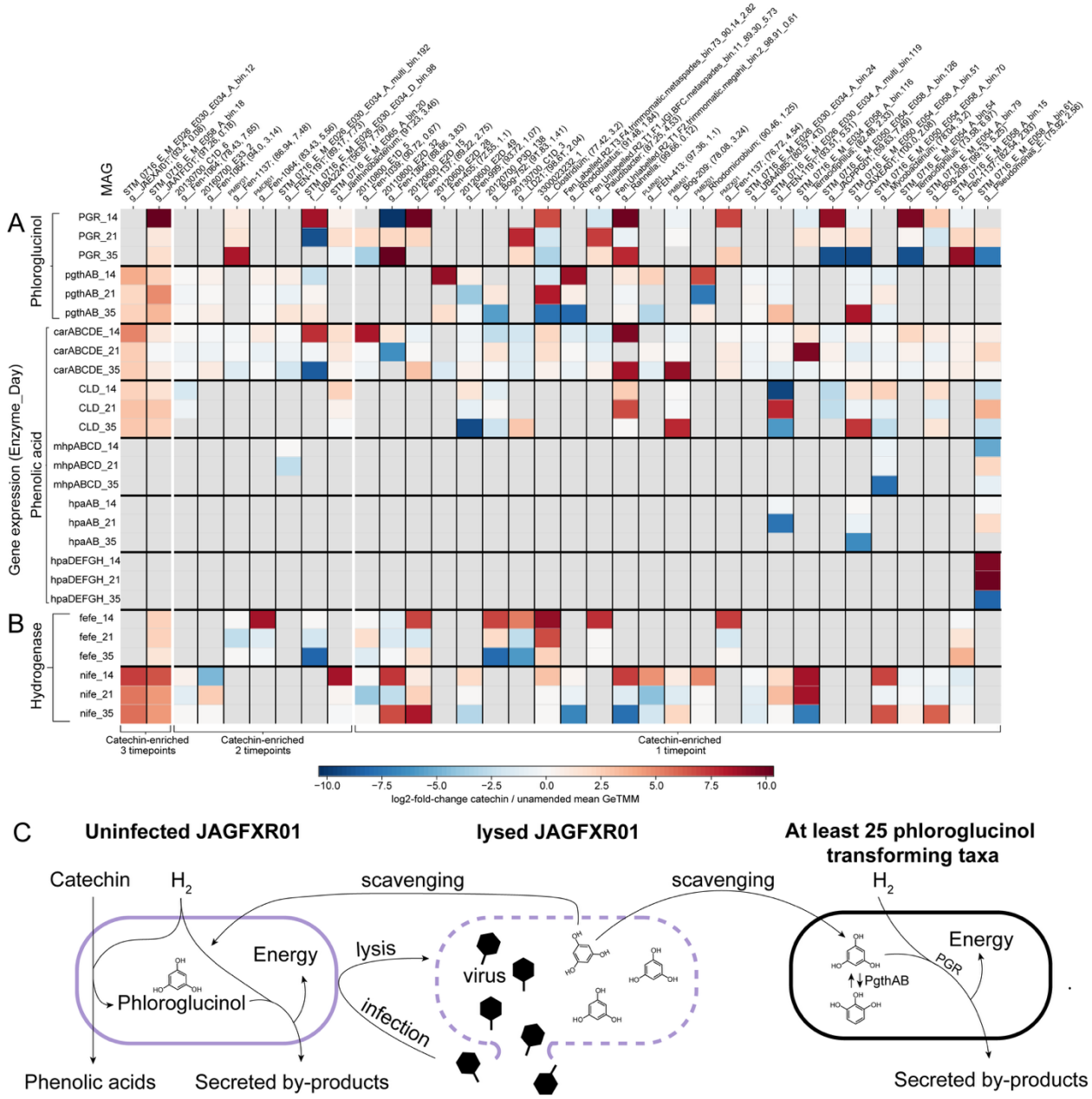
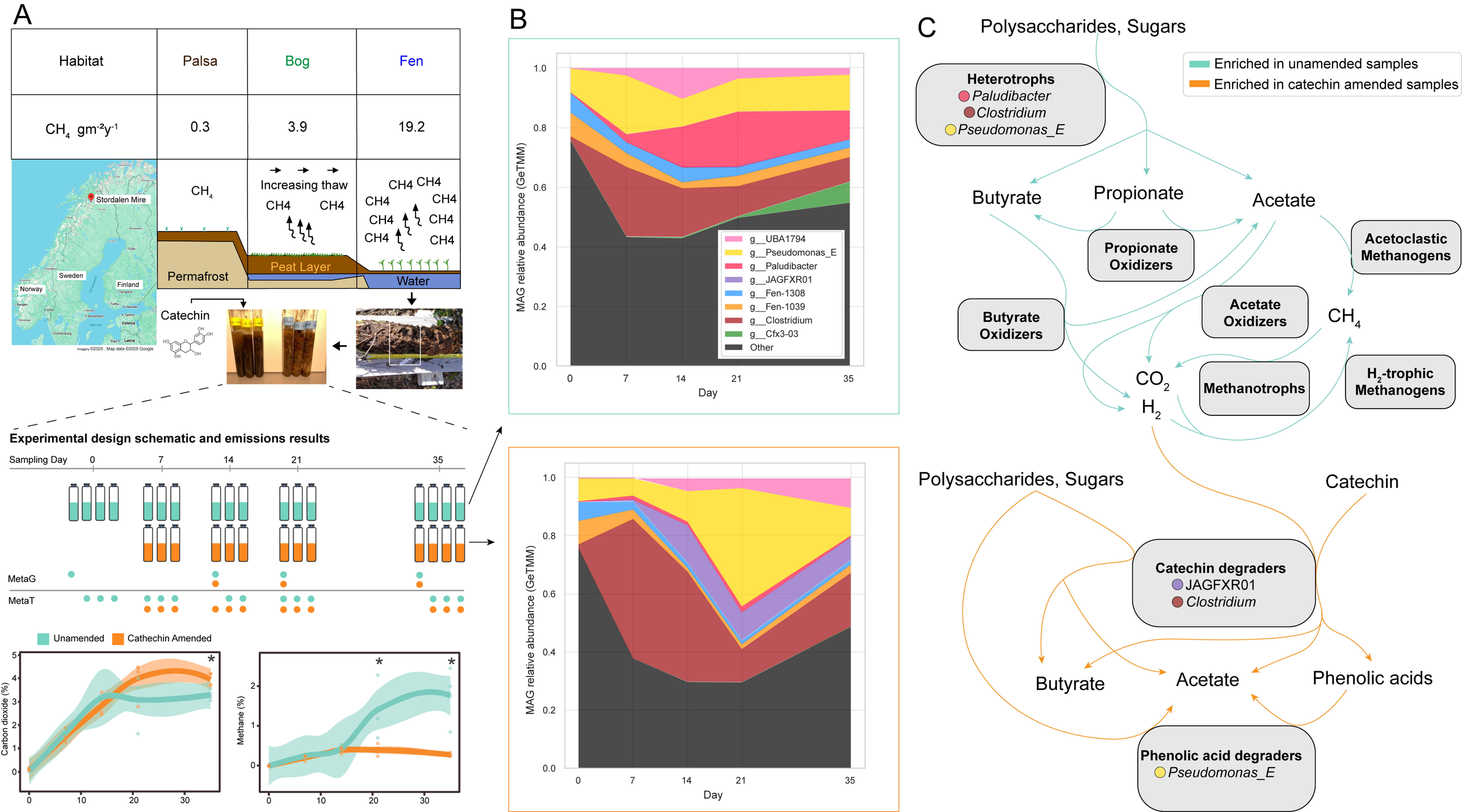
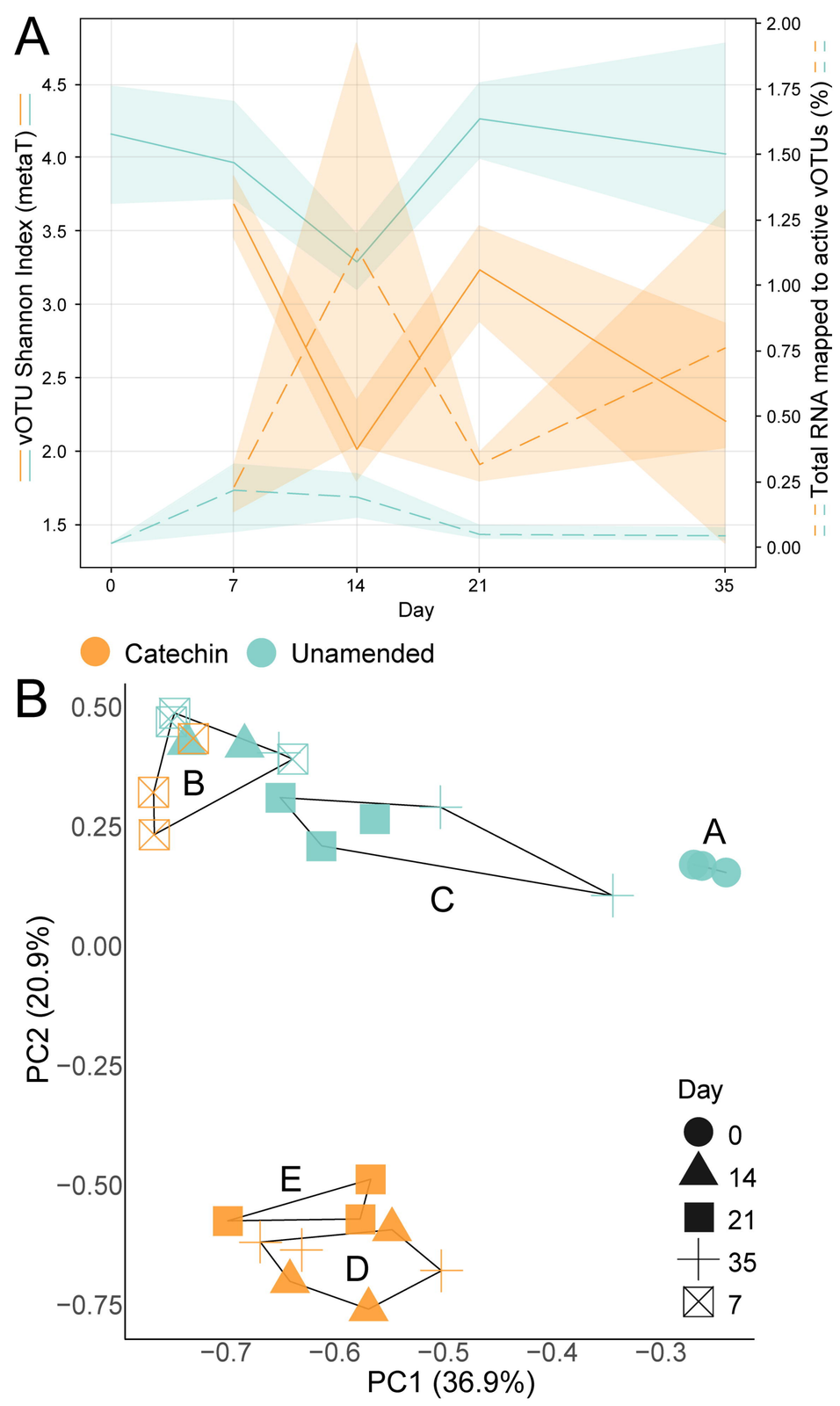
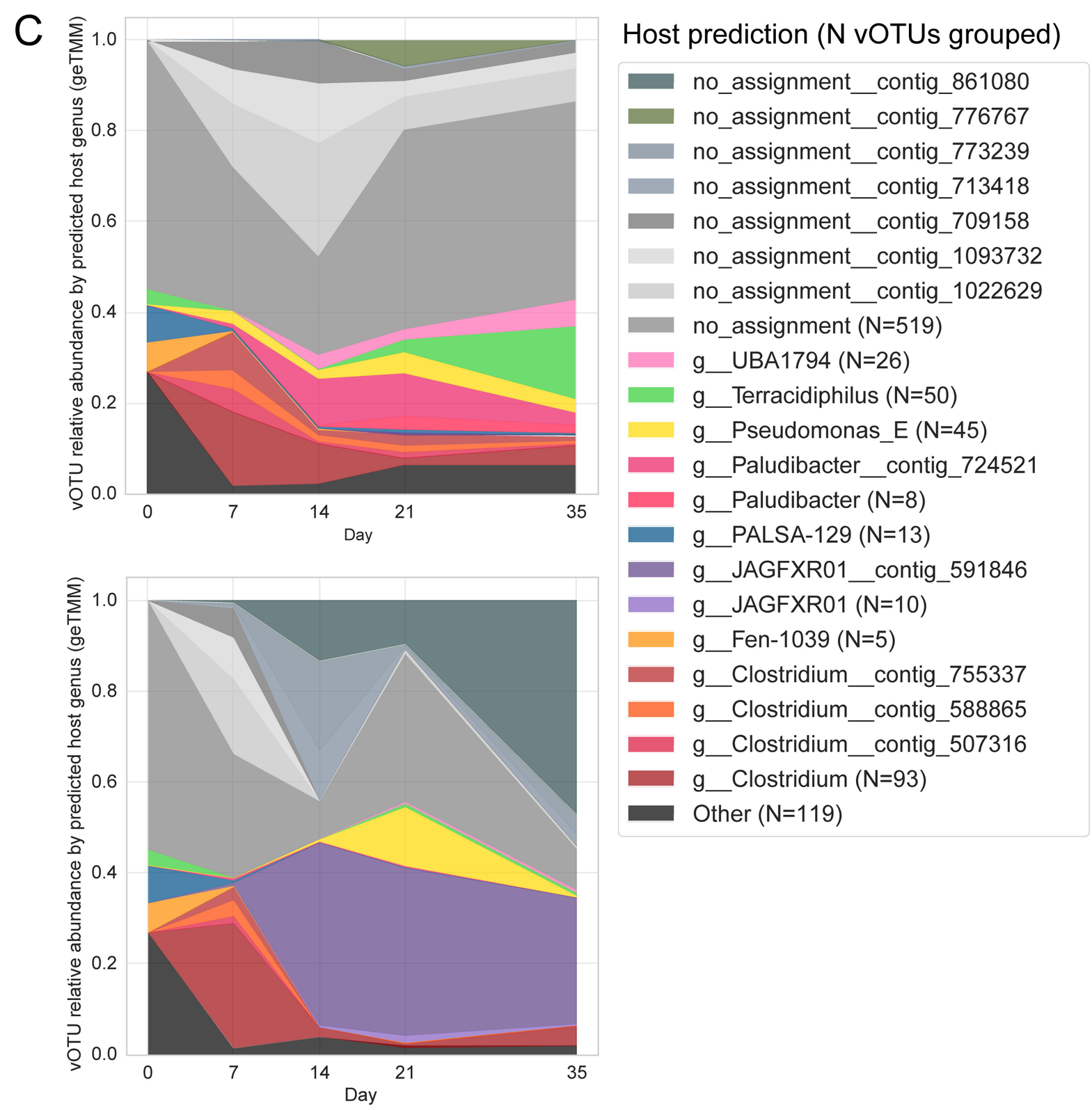


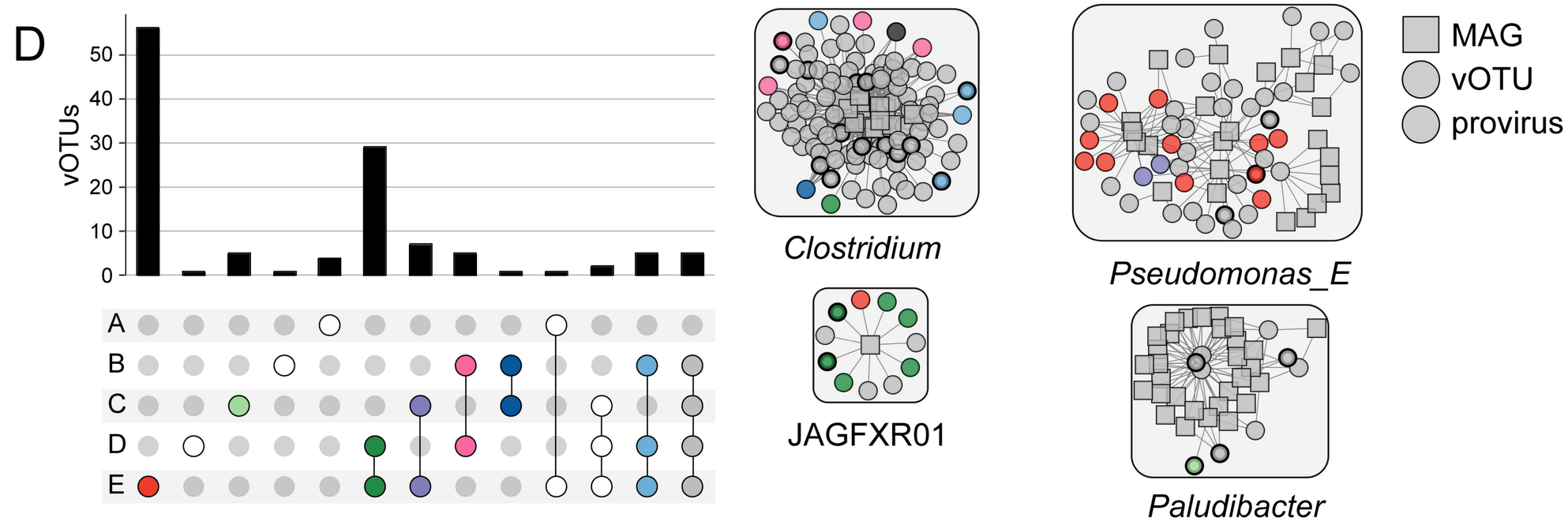

**Figure 5. Proposed mechanism by which JAGFXR01-infecting vOTUs facilitate broader microbial community metabolism. (A)** Heatmap of log2-fold-change expression (GeTMM) between catechin-amended and unamended samples of polyphenol metabolism from non-catechin-degraders. **(B)** Hydrogenase gene log2-fold-change expression (GetMM). Rows are gene expression on a specific sample day. Columns are MAGs that did not express genes for catechin degradation. Below each MAG ID is the GTDB genus classification (r214), CheckM2 completeness, and CheckM2 contamination. Related metabolisms (phloroglucinol, phenolic acid, or hydrogenase) are grouped together. Log2-fold-change was computed by adding a pseudocount equal to the minimum magnitude of GeTMM relative abundance (10e-4), taking the mean of sample replicates within a treatment, then dividing the means between treatments that were taken on the same day. The GeTMM values were summed within each sample for genes with multiple subunits (e.g. pgthAB is the sum of pgthA GeTMM expression and pgthB

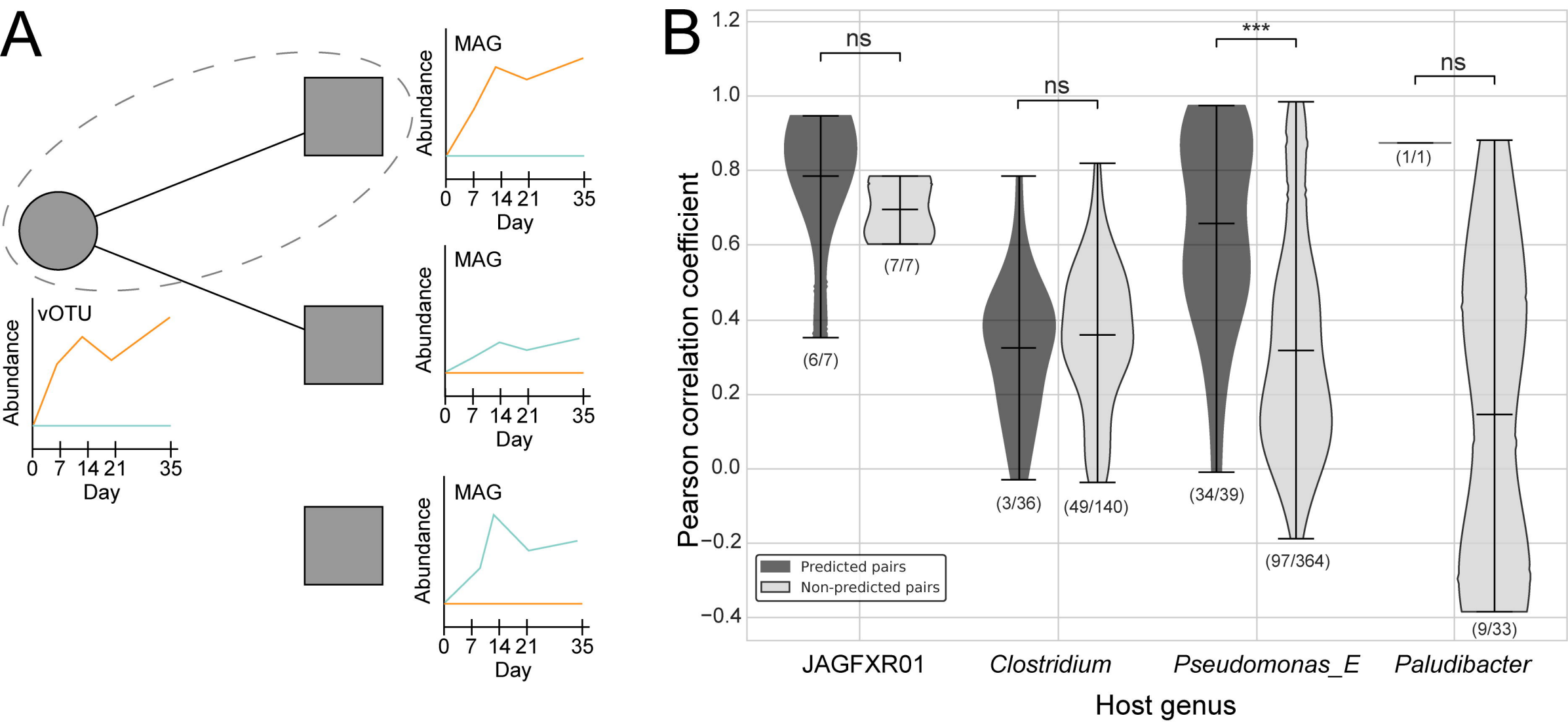
883 GeTMM expression). Warm colors indicate catechin enrichment and cool colors indicate
884 unamended enrichment. Grey indicates no detection in either catechin-amended or unamended
885 samples. **(C)** Lysed JAGFXR01 releases partially metabolized catechin intermediates, including
886 compounds that undergo phloroglucinol transformation, enabling their utilization by downstream
887 community members. JAGFXR01 metabolism is based on previous work [17].
888

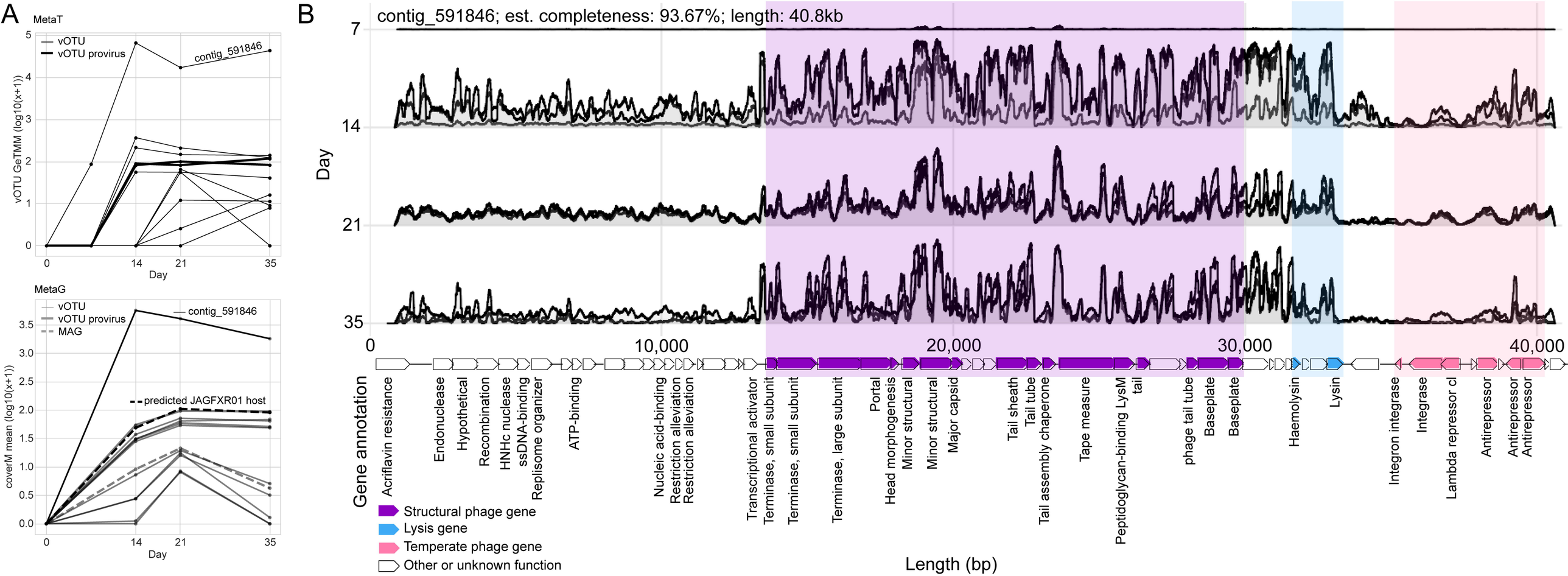

A
Habitat
Palsa
Bog
Fen
CH4 gm-2y-1
0.3
3.9
19.2
Stordalen Mire
Sweden
Norway
Finland
CH4
Increasing thaw
Permafrost
Peat Layer
Water
Catechin
Experimental design schematic and emissions results
Sampling Day
0
7
14
21
35
MetaG
MetaT
Unamended
Cathechin Amended
Carbon dioxide (%)
Methane (%)
B
MAG relative abundance (GeTMM)
Day
g__UBA1794
g__Pseudomonas_E
g__Paludibacter
g__JAGFXR01
g__Fen-1308
g__Fen-1039
g__Clostridium
g__Cfx3-03
Other
C
Polysaccharides, Sugars
Enriched in unamended samples
Enriched in catechin amended samples
Heterotrophs
Paludibacter
Clostridium
Pseudomonas_E
Butyrate
Propionate
Acetate
Propionate Oxidizers
Acetoclastic Methanogens
Butyrate Oxidizers
Acetate Oxidizers
CH4
CO2
H2
Methanotrophs
H2-trophic Methanogens
Polysaccharides, Sugars
Catechin
Catechin degraders
JAGFXR01
Clostridium
Butyrate
Acetate
Phenolic acids
Phenolic acid degraders
Pseudomonas_E

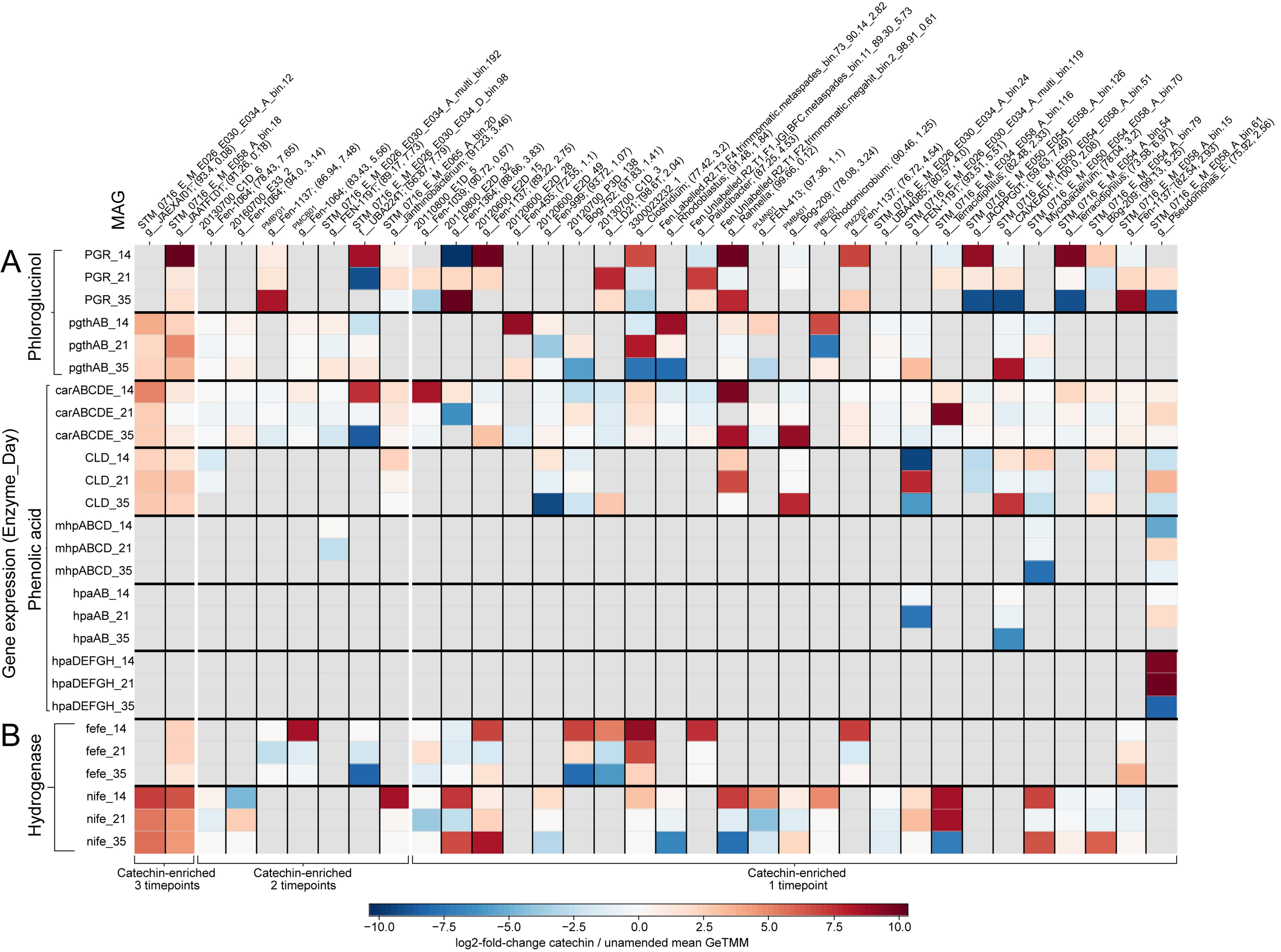

A
vOTU Shannon Index (metaT)
Total RNA mapped to active vOTUs (%)
Day
Catechin
Unamended
B
PC2 (20.9%)
PC1 (36.9%)
Day
0
14
21
35
7
A
B
C
D
E

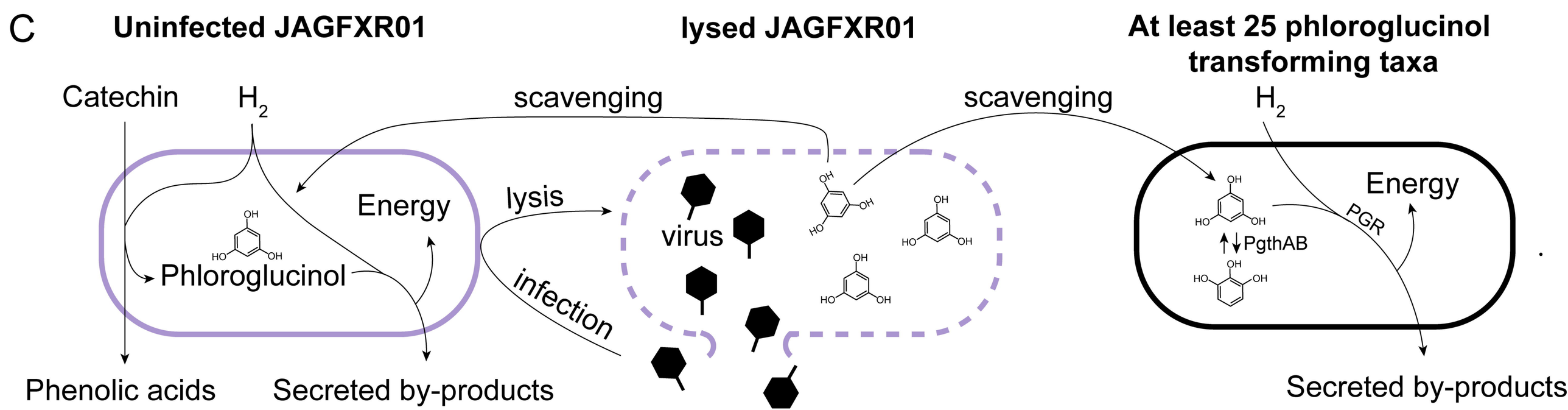

C
vOTU relative abundance by predicted host genus (geTMM)
Day
Host prediction (N vOTUs grouped)
no_assignment__contig_861080
no_assignment__contig_776767
no_assignment__contig_773239
no_assignment__contig_713418
no_assignment__contig_709158
no_assignment__contig_1093732
no_assignment__contig_1022629
no_assignment (N=519)
g__UBA1794 (N=26)
g__Terracidiphilus (N=50)
g__Pseudomonas_E (N=45)
g__Paludibacter__contig_724521
g__Paludibacter (N=8)
g__PALSA-129 (N=13)
g__JAGFXR01__contig_591846
g__JAGFXR01 (N=10)
g__Fen-1039 (N=5)
g__Clostridium__contig_755337
g__Clostridium__contig_588865
g__Clostridium__contig_507316
g__Clostridium (N=93)
Other (N=119)


D
vOTUs
A
B
C
D
E
Clostridium
Pseudomonas_E
JAGFXR01
Paludibacter
MAG
vOTU
provirus


A
vOTU
MAG
MAG
MAG
Abundance
Day
0 7 14 21 35
B
Pearson correlation coefficient
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
ns
ns
***
ns
(6/7)
(7/7)
(3/36)
(49/140)
(34/39)
(97/364)
(1/1)
(9/33)
Predicted pairs
Non-predicted pairs
JAGFXR01
Clostridium
Pseudomonas_E
Paludibacter
Host genus


A
MetaT
vOTU
vOTU provirus
contig_591846
vOTU GeTMM (log10(x+1))
Day
MetaG
vOTU
vOTU provirus
MAG
contig_591846
predicted JAGFXR01 host
coverM mean (log10(x+1))
Day
B
contig_591846; est. completeness: 93.67%; length: 40.8kb
Day
Gene annotation
Acriflavin resistance
Endonuclease
Hypothetical
Recombination
HNHc nuclease
ssDNA-binding
Replisome organizer
ATP-binding
Nucleic acid-binding
Restriction alleviation
Restriction alleviation
Transcriptional activator
Terminase, small subunit
Terminase, small subunit
Terminase, large subunit
Portal
Head morphogenesis
Minor structural
Minor structural
Major capsid
Tail sheath
Tail tube
Tail assembly chaperone
Tape measure
Peptidoglycan-binding LysM
tail
phage tail tube
Baseplate
Baseplate
Haemolysin
Lysin
Integron integrase
Integrase
Lambda repressor cI
Antirepressor
Antirepressor
Antirepressor
Structural phage gene
Lysis gene
Temperate phage gene
Other or unknown function
Length (bp)


MAG
A
Phloroglucinol
PGR_14
PGR_21
PGR_35
pgthAB_14
pgthAB_21
pgthAB_35
Gene expression (Enzyme_Day)
Phenolic acid
carABCDE_14
carABCDE_21
carABCDE_35
CLD_14
CLD_21
CLD_35
mhpABCD_14
mhpABCD_21
mhpABCD_35
hpaAB_14
hpaAB_21
hpaAB_35
hpaDEFGH_14
hpaDEFGH_21
hpaDEFGH_35
B
Hydrogenase
fefe_14
fefe_21
fefe_35
nife_14
nife_21
nife_35
Catechin-enriched 3 timepoints
Catechin-enriched 2 timepoints
Catechin-enriched 1 timepoint
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
log2-fold-change catechin / unamended mean GeTMM


C
Uninfected JAGFXR01
lysed JAGFXR01
At least 25 phloroglucinol transforming taxa
Catechin
$H_2$
scavenging
scavenging
$H_2$
Energy
lysis
virus
infection
Phloroglucinol
Energy
PgthAB
PGR
Phenolic acids
Secreted by-products
Secreted by-products