Supporting Information (SI)

Monier et al. 10.1073/pnas.1708097114

SI Tables

Table S1. List of forward (F) and reverse (R) PCR primers used in the present study.

Primer name	Sequence (5′ ↦ 3′)	PCR product length (bp)	Reference
vAmt-F4	CGCTCGGAATATCCTGGTCC	2370	This study
vAmt-R4	AGAAACCCAAACTTCGGCCA		
OtV6-full-F1	AAGGACCAAAAGTGCCCCAA	2489	This study
OtV6-full-R1	TTTGCGATGTTGTGTGCTGG		
sh-Amt-vir-F	TTCGCGTTACCTACCGAGTG	1757	This study
sh-Amt-vir-R	TACCCTCTGACGAGTGCAGA		
RT-vAmt-F1	CCACCACCGTAGAAGACACC	642	This study
RT-vAmt-R1	TGGTCCAAAGCTTCTAGGCG		
RT-vAmt-F3	TACACCTAGGGCACCACAGA	606	This study
RT-vAmt-R3	CATTGGGGTTGGTCCACTGA		
Euk82F	GAAACTGCGAATGGCTC	1030	López-García *et al.*, 2003 [1]
nonMet-R	TTTAAGTTTCAGCCTTGCG		Bower *et al.*, 2003 [2]
vAmt-attF	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGATGGACATCCAAACTTTGT	1378	This study
vAmt-attR	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAGCCATTTCGTAAGCAG		

Table S2. Temporal changes in cell abundances, ammonium concentrations and methylammonium uptake rates of *O. tauri* cultures.

Measurement†	hpi‡	Control mean ±s.e.	Infected mean ±s.e.	Control fold change§	Infected fold change§	Significance*
FCM cell counts	0	4.6 ±0.6	4.1 ±0.6	0	0	-
	2	5.1 ±0.5	4.2 ±0.6	0.17 ±0.07	0.02 ±0.02	*T*=2.083; *P*=0.105
	4	4.7 ±0.2	4.2 ±0.6	0.06 ±0.1	0.02 ±0.005	*T*=0.290; *P*=0.786
	6	5.2 ±0.4	4.7 ±0.7	0.21 ±0.08	0.18 ±0.0005	*T*=0.339; *P*=0.751
	8	5.2 ±1	4.9 ±0.7	0.16 ±0.15	0.25 ±0.06	*T*=-0.546; *P*=0.614
	12	6.7 ±0.7	4.3 ±0.6	0.55 ±0.05	0.07 ±0.04	***T*=7.029; *P*=0.002**
	16	7.2 ±0.8	4.9 ±0.6	0.65 ±0.05	0.23 ±0.009	***T*=8.360; *P*=0.001**
[NH_4^+] medium	0	576 ±21.6	590 ±6.65	0	0	-
	2	545 ±5.29	569 ±22.11	-0.07 ±0.05	-0.05 ±0.06	*T*=-0.281; *P*=0.792
	4	535 ±5.56	560.66 ±34.10	-0.10 ±0.05	-0.07 ±0.09	*T*=-0.230; *P*=0.829
	6	535 ±8.38	538.33 ±24.88	-0.10 ±0.06	-0.13 ±0.06	*T*=0.308; *P*=0.772
	8	544 ±32.00	514.66 ±22.16	-0.08 ±0.11	-0.19 ±0.06	*T*=0.840; *P*=0.448
	12	561.66 ±25.64	540 ±7.53	-0.03 ±0.120	-0.13 ±0.06	*T*=0.671; *P*=0.538
	16	504.66 ±10.89	485.66 ±18.22	0.19 ±0.07	-0.23 ±0.03	*T*=1.043; *P*=0.355
Methylammonium uptake	0	0.063 ±0.016	0.03 ±0.004	0	0	-
	2	0.042 ±0.011	0.044 ±0.013	-0.57 ±0.06	0.35 ±0.27	***T*=-3.273; *P*=0.03**
	4	0.053 ±0.021	0.049 ±0.012	-0.39 ±0.2	0.55 ±0.24	*T*=-2.997; *P*=0.4
	6	0.083 ±0.021	0.067 ±0.008	0.39 ±0.02	1.08 ±0.11	***T*=-6.282; *P*=0.003**
	8	0.0309 ±0.003	0.026 ±0.005	-0.98 ±0.21	-0.30 ±0.16	***T*=-2.439; *P*=0.007**
	12	0.087 ±0.021	0.089 ±0.026	0.46 ±0.08	1.38 ±0.3	***T*=-2.919; *P*=0.043**
	16	0.041 ±0.006	0.033 ±0.005	-0.58 ±0.23	0.05 ±0.04	*T*=-2.394; *P*=0.074

†FCM counts in 10^7.cell.mL^{-1}; [NH_4^+] in fluorescent units (1 µM NH_4^+: 1086 ±12.4 s.e.); [^{14}C]-methylammonium uptake rate in nmol.mg^{-1}.min^{-1}.
‡Hour post infection. §Log2 fold changes (0 hpi based). *Two-sample t-tests (df= 4) on log2 fold changes between control/infected triplicates.

Table S3. OtV6 particle abundances in infected *O. tauri* cultures.

hpi†	mean ±s.e.†	Fold change§	Significance*
0	1.68 ±0.04	0	-
2	1.67 ±0.05	-0.01 ±0.01	*T*=-0.768; *P*=0.522
4	1.64 ±0.08	-0.04 ±0.03	*T*=0.674; *P*=0.537
6	1.39 ±0.11	-0.27 ±0.07	***T*=2.925; *P*=0.043**
8	1.36 ±0.05	-0.30 ±0.03	*T*=0.307; *P*=0.773
12	1.58 ±0.09	-0.09 ±0.05	***T*=-3.354; *P*=0.028**
16	2.60 ±0.18	0.62 ±0.09	***T*=-6.725; *P*=0.002**

†Hour post infection. ‡Based on FCM viral particle counts (10^7.particle.mL^{-1}); no particles detected in non-infected, control *O. tauri* cultures at the start of the experiment (timepoint 0 h). §Log2 fold changes (0 hpi based). *Two-sample t-tests (df=4) on triplicate log2 fold changes between current/previous timepoints (one-sample t-test for 2 hpi).

Table S4. List of qPCR forward (F), reverse (R) primers and probes (P) designed for this study.

Primer name†	UniProtKB accession	Sequence (5′ ↦ 3′)	PCR product length (bp)	Efficiency (%)
Ot-betaTubulin-F		CAACGTTAAGAGCTCGGTGT		
Ot-betaTubulin-R	A0A090LXT7	CCATCTCATCCATACCCTCA	179	98.27
Ot-betaTubulin-P		TAAGATGTCCGCCACCTTTG		
Ot-amt-F		ACCAAGTTCGAGGCTTACCT		
Ot-amt-R	A0A096PA30	CAGTCATGTGGACGATACCA	196	100.87
Ot-amt-P		TCCTCTGCGCATTCGTCTAC		
OtV6-polB-F		AGAGGAGAATGGGGTACCAG		
OtV6-polB-R	H8ZJQ3	GACGAGTTCGTCCTTCTTCA	193	96.70
OtV6-polB-P		AAACCAAAGAAGCCTCCGAA		
OtV6-mcp-F		CTCACCGACTTCAAGCTCAT		
OtV6-mcp-R	H8ZJG5	CGCGAAGGAGTAGGAGTACA	135	99.13
OtV6-mcp-P		AGGTCCAGGCCTACAACCAC		
OtV6-amt-F		GAGCAACCGGCAGTAATAGA		
OtV6-amt-R	H8ZJB2	AAGGTTATCTCACGCACAGC	158	101.30
OtV6-amt-P		GAGAGCCATCGCTGTGAGAC		

SI Figures

Fig. S1. Ammonium transporter domain architecture, gene structure and C-terminus alignment of vAmt and Amt-Euk homologs.

Fig. S2. Maximum-likelihood phylogeny of green algal viruses based on DNA polymerase B protein sequences.

Fig. S3. Score distribution of BLASTP pairwise searches of all ORFs from *Ostreococcus* and *Micromonas* viruses.

Fig. S4. Maximum-likelihood phylogenies of vAmt flanking genes.

Fig. S5. RT-DNA gel plot of *O. tauri* infected cells, RT-PCR-targeting OtV6 vAmt.

Fig. S6. *O. tauri* Amt1.1 and vAmt sequence and structure comparisons.

Fig. S7. *O. tauri* Amt1.1 and vAmt transmembrane domain predictions.

Fig. S8. Functional complementation experiments of ammonium or urea uptake defective yeasts for various nitrogen sources.

Fig. S9. Respiration curves from the OmniLog® Phenotype MicroArray system.

Fig. S10. Ammonium concentrations in growth media.

Fig. S11. Amt/Mep/Rh superfamily phylogenetic tree.

Fig. S12. vAmt evolutionary relationships with Amt-Euk homologs.

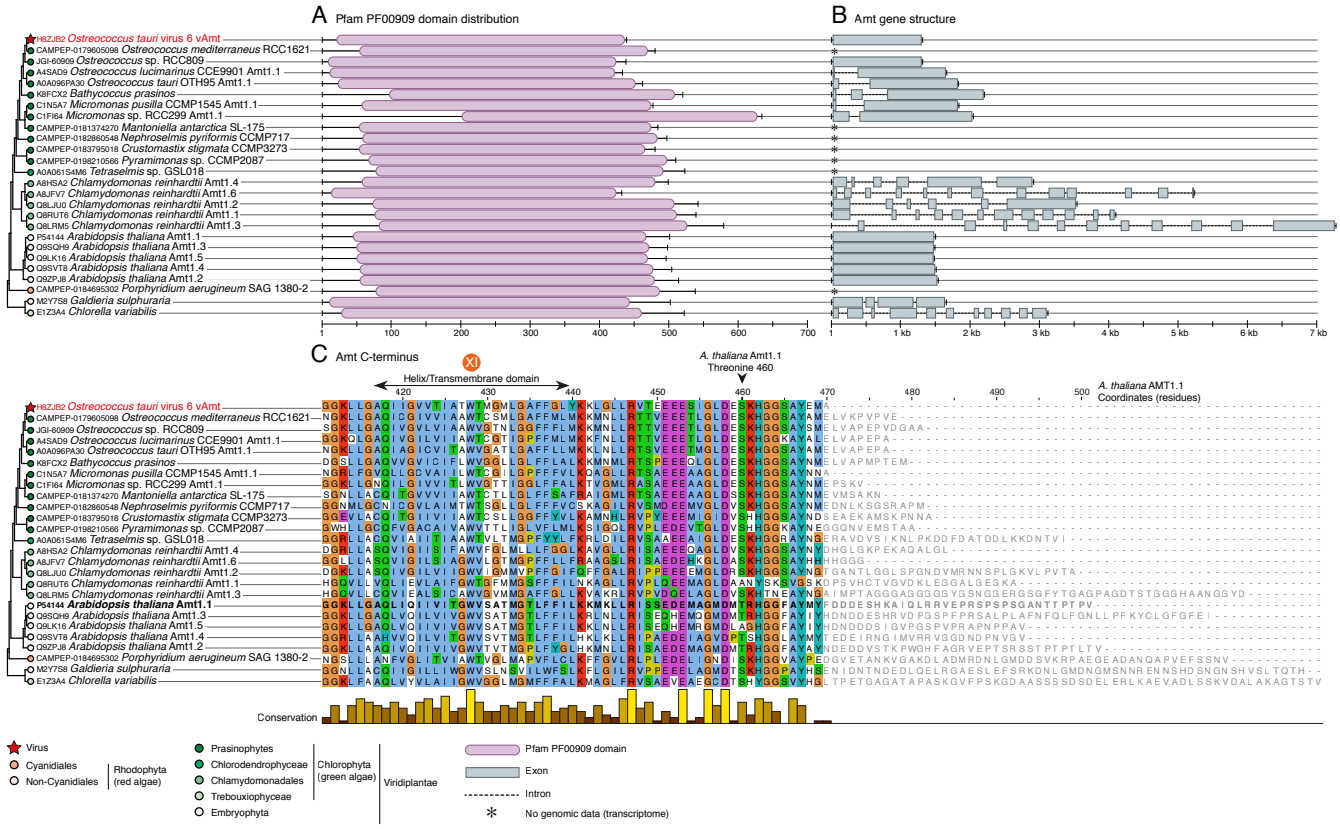

Fig. S1. Ammonium transporter domain architecture (*A*), gene structure (*B*) and C-terminus alignment of vAmt and Amt-Euk (Amt1) homologs (*C*). The eukaryotic sequences shown here were selected based on phylogenetic analysis (Fig. 6B and Fig. S11) to identify homologs of taxa that branch in key positions relative to vAmt. For each sequence, the taxonomic information is represented as a colored circle with shades of green or red color for Chlorophytes (green algae) and Rhodophyta (red algae), respectively. The vAmt is displayed in red characters and a red star indicates its position. (*A*) Pfam PF00909 domain distribution along the protein sequences (lower scale indicates amino acid residue position); domains are represented by pink oblongs. (*B*) Corresponding gene structure (lower scale indicates nucleotide position in kb); exons are represented by grey boxes and introns are dashed lines. Asterisks indicate sequence data originating from transcriptome, for which gene structure information is lacking. Note that the vAmt gene is intron-less; all other homologs shown here, with available gene structure information, possess at least one intron, with the exception of the homolog from *Ostreococcus* sp. RCC809 and *Arabidopsis*. (*C*) Alignment of vAmt and Amt-Euk C-terminal domains showing that the vAmt protein is shorter than its most similar homologs in eukaryotes, with the exception of *Micromonas pusilla* CCMP1545. The upper scale corresponds to the amino acid residue position of *Arabidopsis thaliana* Amt1.1, which is shown in bold characters. A black arrowhead indicates the threonine 460 of *A. thaliana* Amt1.1. Phosphorylation of T460 of *A. thaliana* Amt1.1 may act as an allosteric switch to limit ammonium transport [3]; T460 is substituted by a serine in green and red algae as well as in the vAmt sequence; note that in eukaryotes, serine is also a phosphorylation site. Position of the transmembrane helix XI is indicated on the upper scale. The histogram below the sequence alignment represents the amino acid residue conservation.

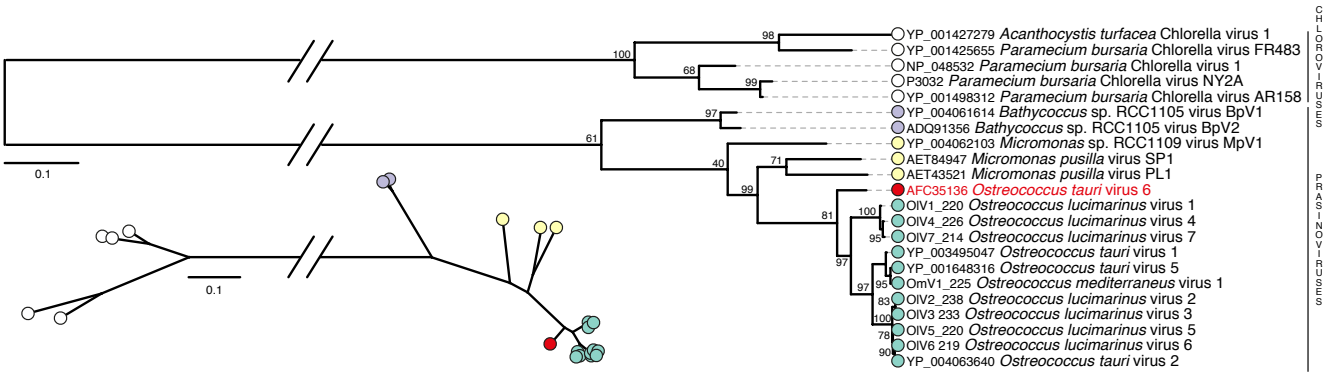

Fig. S2. Maximum-likelihood phylogenetic tree of green algal viruses inferred from a DNA polymerase B alignment (915 sites) under the LG+G+I+F model, as implemented in RAxML [4]. The unrooted version of this tree is presented below the midpoint-rooted version. Viral taxonomic information is color coded as in Fig. 1. Accession numbers displayed in front of viral taxonomic names are RefSeq identifiers. Branch node supports were computed from 1000 non-parametric bootstrap replicates. The scale bar represents the number of estimated substitutions per site. The branch connecting the prasinoviruses to the *Chlorella* viruses (chloroviruses) was truncated to allow tree display.

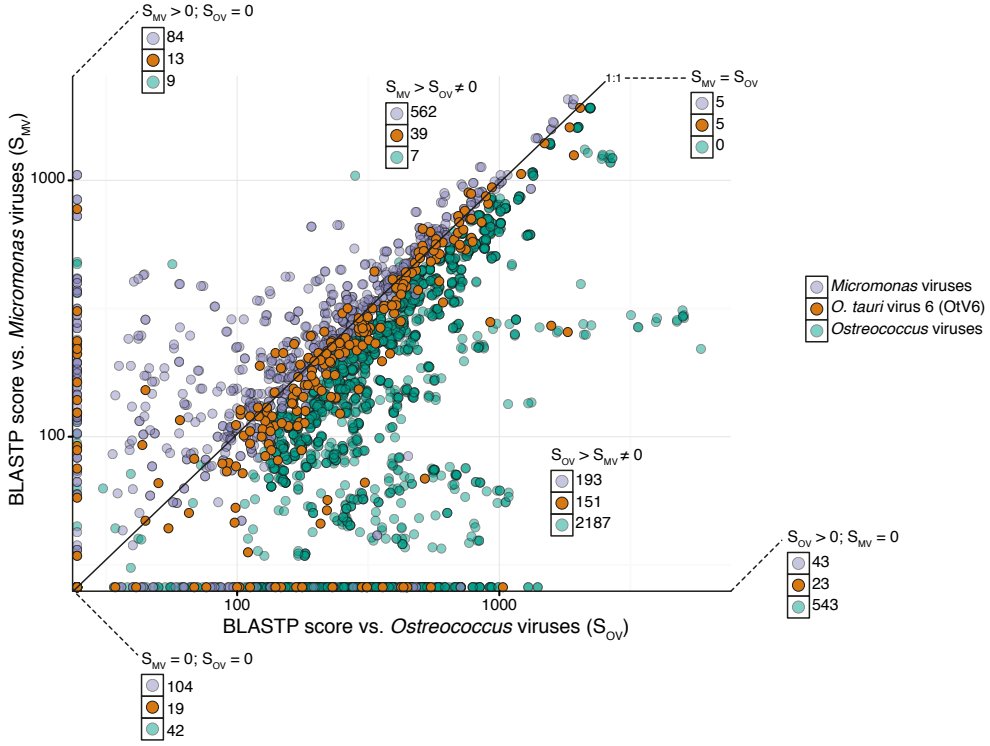

Fig. S3. Score distribution of BLASTP pairwise searches of all ORFs from *Ostreococcus* and *Micromonas* viruses. Dots represent a comparison of BLASTP scores of each ORF between their best BLASTP hits with *Ostreococcus* (x-axis; S_{OV}) and *Micromonas* (y-axis; S_{MV}) viruses, excluding self-hits. Orange, purple and green dots represent ORFs from OtV6 (total: 250 ORFs), *Micromonas* viruses (991 ORFs) and *Ostreococcus* viruses (2788 ORFs, without OtV6), respectively. BLASTP score distributions were categorized as followed: closer to *Micromonas* viruses (i.e., BLASTP score against a *Micromonas* virus sequence is greater than against a sequence from an *Ostreococcus* virus, denoted as $S_{MV} > S_{OV}$), unique to *Micromonas* viruses ($S_{MV} > 0$ and $S_{OV} = 0$), closer to *Ostreococcus* viruses ($S_{OV} > S_{MV}$), unique to *Ostreococcus* viruses ($S_{OV} > 0$ and $S_{MV} = 0$), equidistant between *Micromonas* viruses and *Ostreococcus* viruses ($S_{MV} = S_{OV}$) and unique to OtV6 ($S_{MV} = 0$ and $S_{OV} = 0$). BLASTP score distribution categories are indicated directly on the graph, along with the number of ORFs falling within each category.

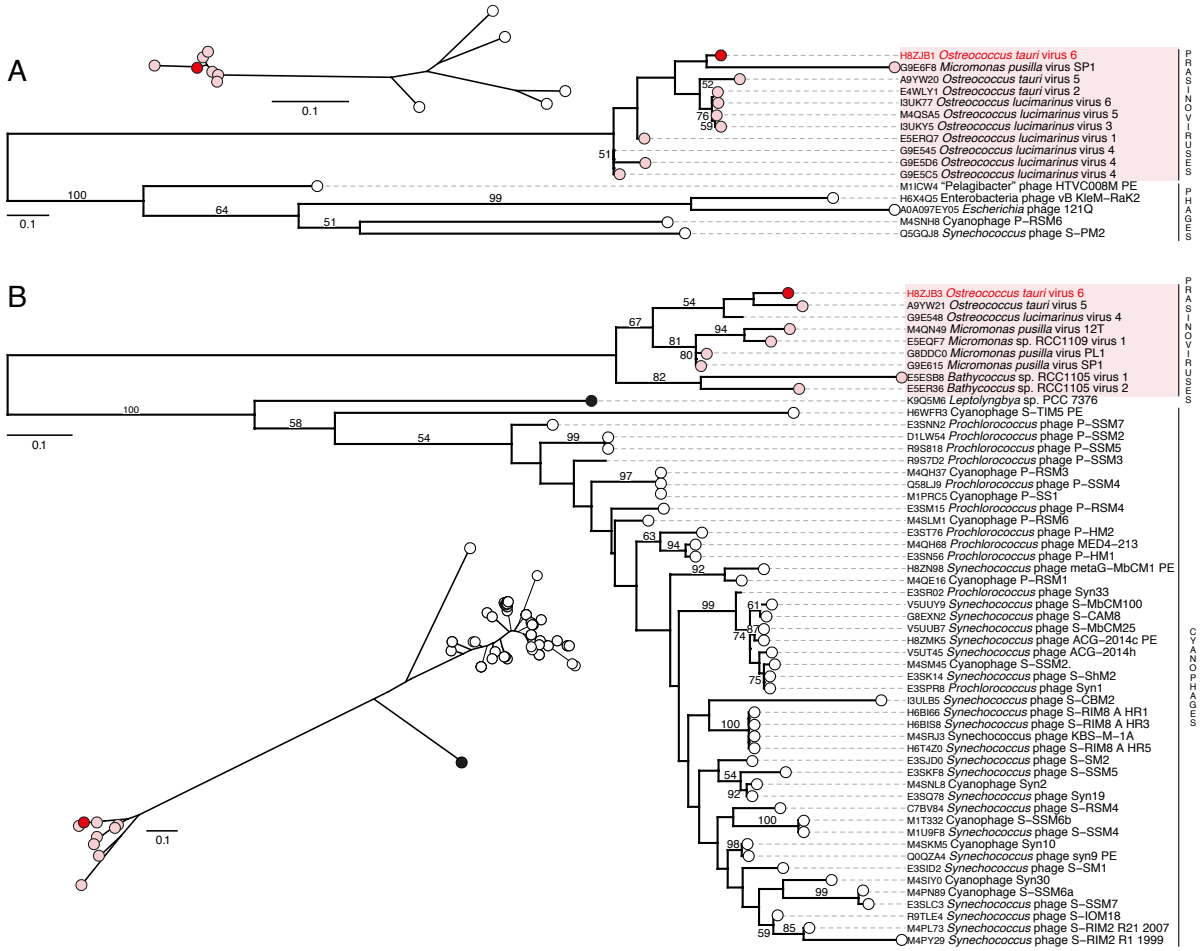

Fig. S4. vAmt flanking gene phylogenies. (*A*) OtV6_114c and (*B*) OtV6_116 maximum-likelihood phylogenetic trees, reconstructed using RAxML with LG+G+I and WAG+G for OtV6_114c and OtV6_116 phylogenies, respectively. For both reconstructions, protein sequence homologs were retrieved using BLASTP searches against UniProtKB; homologous sequences were then aligned using MAFFT with the E-INS-i iterative refinement algorithm [5]. Alignment sites were selected using the gap distribution selection mode implemented in trimAl [6], and sites composed of more than 50% of gaps were removed from the alignments. Local supports were computed from 100 non-parametric bootstrap replicates; only values >50% are displayed. Unrooted versions of the trees are presented below the midpoint-rooted versions. For each sequence, the taxonomic information is represented as a colored circle: light red, white and black circles represent eukaryotic viruses (all from prasinoviruses in both phylogenies), bacteriophages and cellular organisms (only one sequence in the OtV6_116 phylogeny, originating from the bacteria *Leptolyngbya*), respectively. OtV6_114c and OtV6_116 sequences are displayed in red and marked by a red circle. For each tree, the prasinovirus clade is highlighted by a pink rectangle. The scale bar represents the number of estimated substitutions per site.

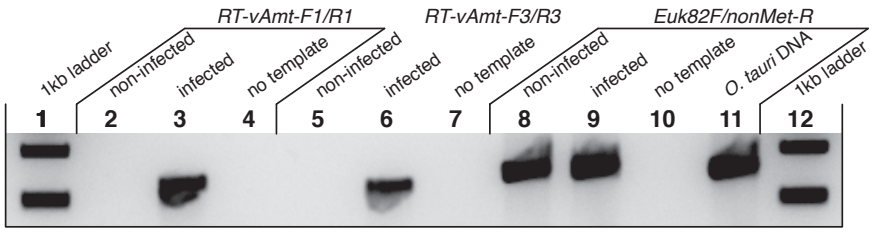

Fig. S5. vAmt is expressed during OtV6 infection cycle of *O. tauri*. RT-DNA gel plot of infected and non-infected *O. tauri* cells, sampled 12 h post OtV6 inoculation, targeting the vAmt gene built from amplified RT-PCR. The vAmt PCR amplifications were conducted using two distinct sets of primers (RT-vAmt-F1/R1 and F3/R3; see Table S1), as well as experimental controls (no template, Euk82F/nonMet-R primer set targeting the 18S rRNA gene; see Table S1).

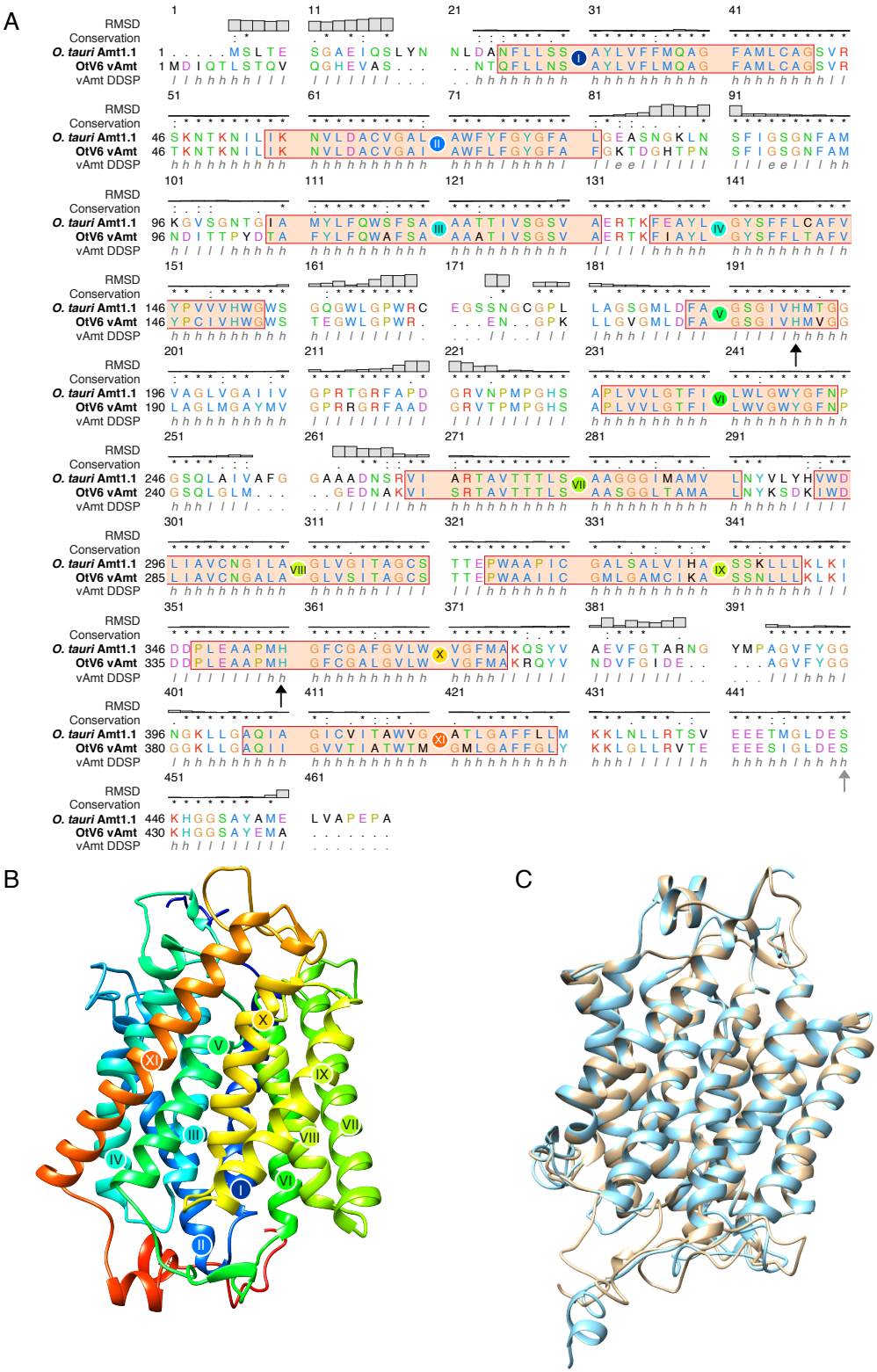

Fig. S6. *O. tauri* Amt1.1 and vAmt sequence and structure comparisons. (*A*) *O. tauri* and OtV6 vAmt sequence comparison. Both protein sequences were aligned using EMBOSS Needle global pairwise alignment [7]. Eleven transmembrane domains (noted I to XI with circle color code as in Fig. 2B and 2C), predicted with TMHMM v2 [8], are highlighted by light orange rectangles. Amino acid conservations, in Clustal color code format [9], are indicated above the pairwise alignment along with RMSD distance levels, which are based on the superimposition of predicted vAmt and *O. tauri* Amt1.1 structures. The vAmt secondary structure prediction, computed using the DDSP algorithm implemented in DaliLite [10], is displayed below the pairwise alignment with the alpha helices noted 'h' and matching the 11 transmembrane domains. Conserved histidine residues in helices V and X are indicated by black arrows. A grey arrow indicates the serine residue conserved with the threonine 460 in *Arabidopsis thaliana* Amt1.1 (see Fig. S1C). (*B*) *O. tauri* Amt1.1 3D protein structure prediction. The predicted protein structure is composed of 11 alpha-helices corresponding to the transmembrane domains indicated by colored circles and Roman numerals. (*C*) Superimposition of the vAmt and *O. tauri* Amt1.1 3D protein structure models. The viral and host Amt structure models are colored in blue and brown, respectively.

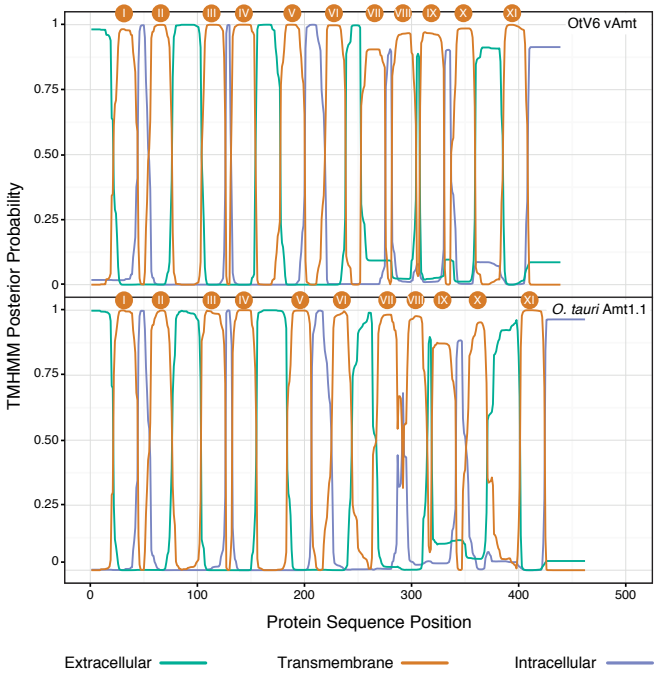

Fig. S7. *O. tauri* Amt1.1 and vAmt transmembrane domain predictions. Transmembrane domains were predicted using TMHMM v2 [8] and the 11 transmembrane domains are noted I to XI with circle color code as in Fig. 2B, C and Fig. S6A, B. The x-axis corresponds to the protein sequence position and the y-axis to the TMHMM posterior probability . Extracellular, transmembrane and intracellular (cytosolic) domains of each protein are represented in green, orange and purple, respectively. Both proteins have an extracellular N-terminus and cytosolic C-terminus, a hallmark of the protein members of the Amt/Mep/Rh superfamily.

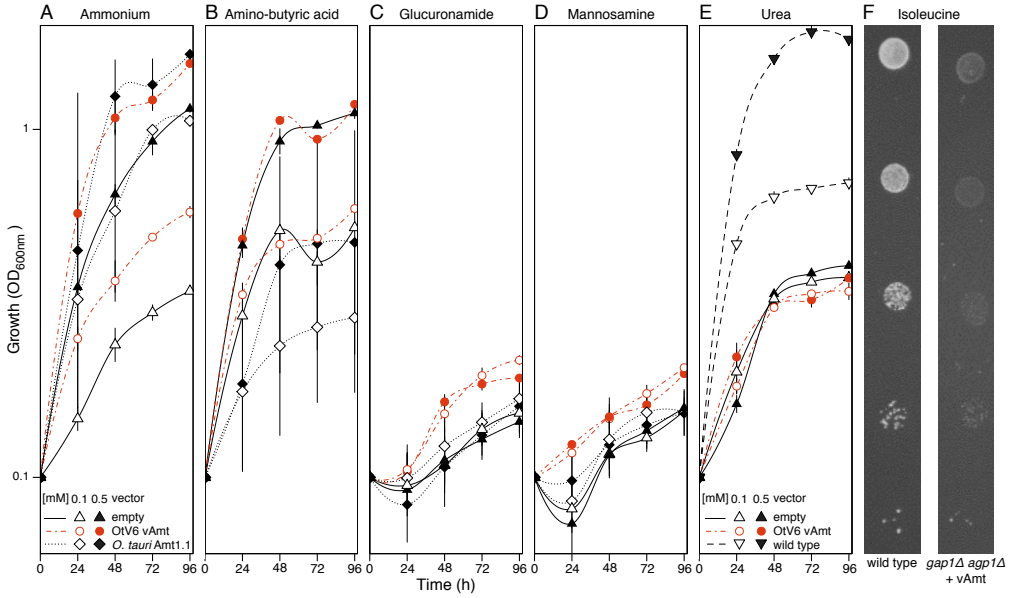

Fig. S8. Culture optical density (OD_{600nm}) of (*A* – *D*) Ammonium-uptake deficient yeast 31019b cultures transformed either with an empty vector (pAG416 GPD; triangles), a vAmt containing vector (pAG416 GPD vAmt; circles) or an *O. tauri* Amt1.1 containing vector (pAG416 GPD Amt1.1; diamonds) grown in media for which the sole nitrogen source was either ammonium sulfate (*A*) , D,L-a-amino-butyric (*B*) acid, glucuronamide (*C*) or D-mannosamine (*D*) at 0.1 and 0.5 mM (black or white filled symbols, respectively). (*E*) OD_{600nm} of an urea (CH_4N_2O) uptake defective yeast mutant, strain YNVW1 (*dur3*Δ), transformed either with pAG416 GPD (triangles) or pAG416 GPD vAmt (circles), and wild type strain Σ23346c transformed with pAG416 GPD (nabla symbol), in two CH_4N_2O concentrations (0.1 and 0.5 mM; black or white filled symbols, respectively). Dots represent OD_{600nm} means of each culture triplicate and error bars represent standard errors. Lines are local polynomial regression fits (loess). (*F*) Yeast spot assay of growth phenotypes for wild type strain (23346c) and pAG416 GPD vAmt-transformed mutant 30633c, which bears deletions in the genes encoding amino acid permeases (*gap1*Δ *agp1*Δ; [11]). Cells were ten-fold serially diluted and spotted onto minimal medium containing 1 mM isoleucine as the sole nitrogen source. The gene encoding vAmt did not complement the growth defect of the *gap1*Δ *agp1*Δ mutant.

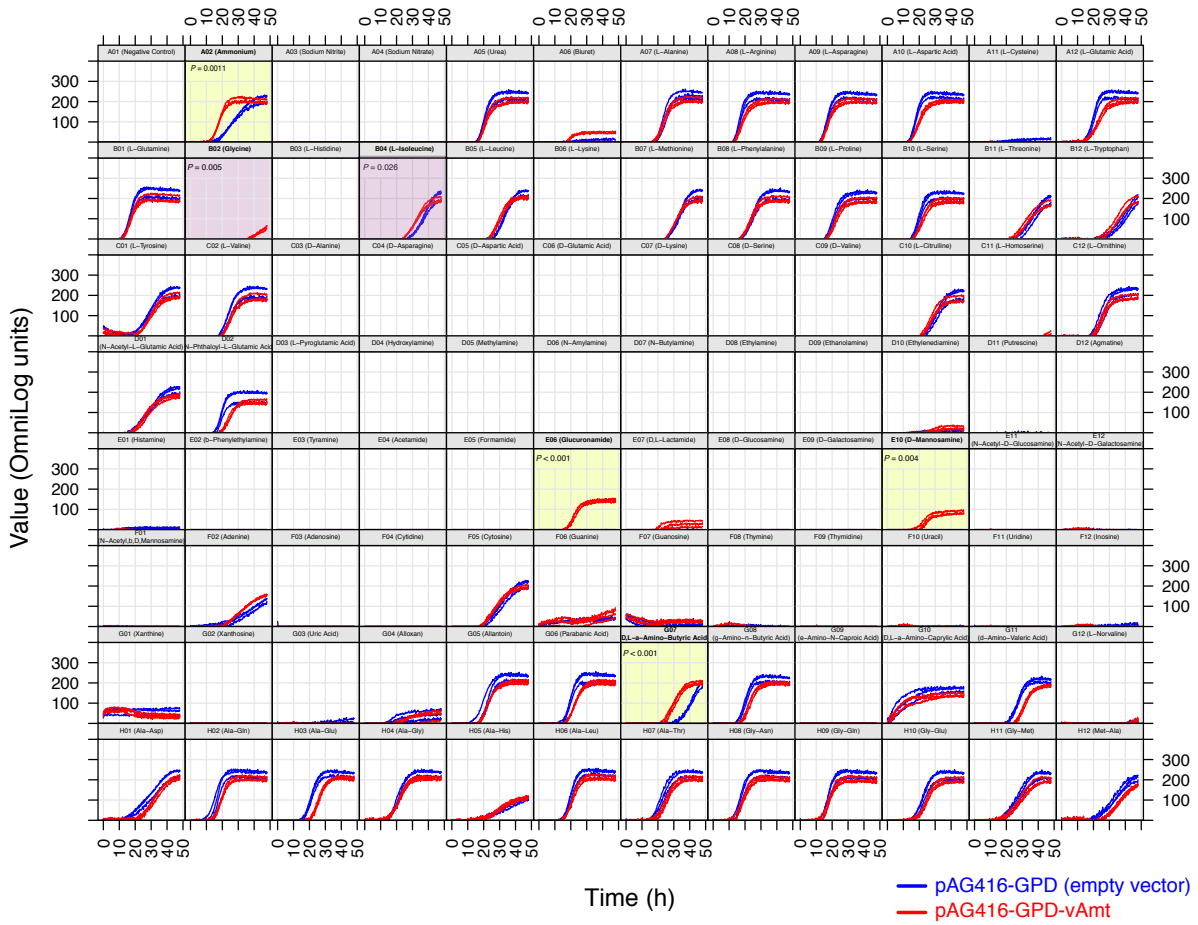

Fig. S9. Respiration curves from the OmniLog® Phenotype MicroArray (PM) system. Ammonium-uptake deficient yeast strain (31019b) cultures, in triplicates, transformed with either pAG416 GPD (empty vector) or pAG416 GPD vAmt were grown on various nitrogen sources from a PM3B MicroPlate™ and have their curves colored in blue and red, respectively. PM curves were generated using the R package *opm* [12]. Yellow shaded plots represent nitrogen sources for which significant AUC (area under curve; see [12]) differences ($P < 0.005$) in PM curves between vAmt/empty-vector transformed culture triplicates were detected, and corresponding *P*-values are displayed. Purple shaded plots represent two amino acid sources, glycine and L-isoleucine producing AUC differences with *P*-values of 0.005 and 0.026, respectively, between vAmt/empty-vector transformed triplicates. No isoleucine phenotype was found for vAmt-transformed yeast mutant (strain 30633c, *gap1*Δ *agp1*Δ; see Fig. S8F).

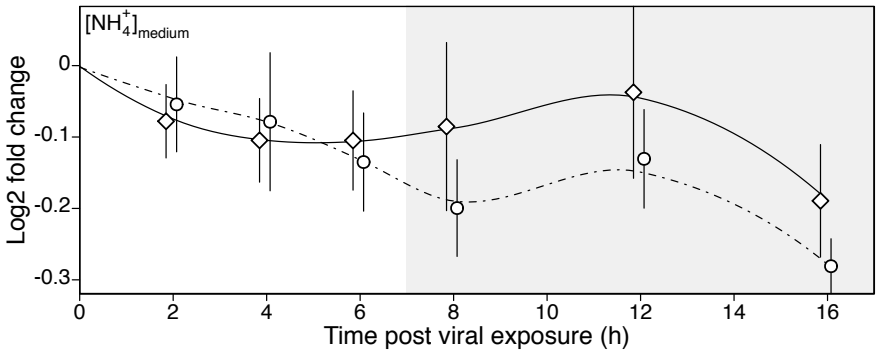

Fig. S10. Temporal dynamics of NH_4^+ concentrations in growth media in OtV6-infected and non-infected *O. tauri* cultures. NH_4^+ concentrations in growth media were determined by fluorometry. Three *O. tauri* cultures were exposed to OtV6 (timepoint 0 h) and three other non-infected cultures were used as controls (same *O. tauri* cultures and experiments than results presented as Fig. 4). Log2 fold changes are based on timepoint 0 h and the plot summarizes the fold change distribution of three cultures for a given timepoint during the experiment; trend lines were estimated by local polynomial regressions (loess). Shaded overlay indicates incubator dark period ('night'). This analysis does not show a significant change in NH_4^+ depletion between infected and non-infected *O. tauri* cultures (see Table S2 for two-sample t-test results), although hints at a trend of increased NH_4^+ uptake in infected cells. To specifically test this we therefore conducted radiolabeled methylammonium uptake assays (see Fig. 4B).

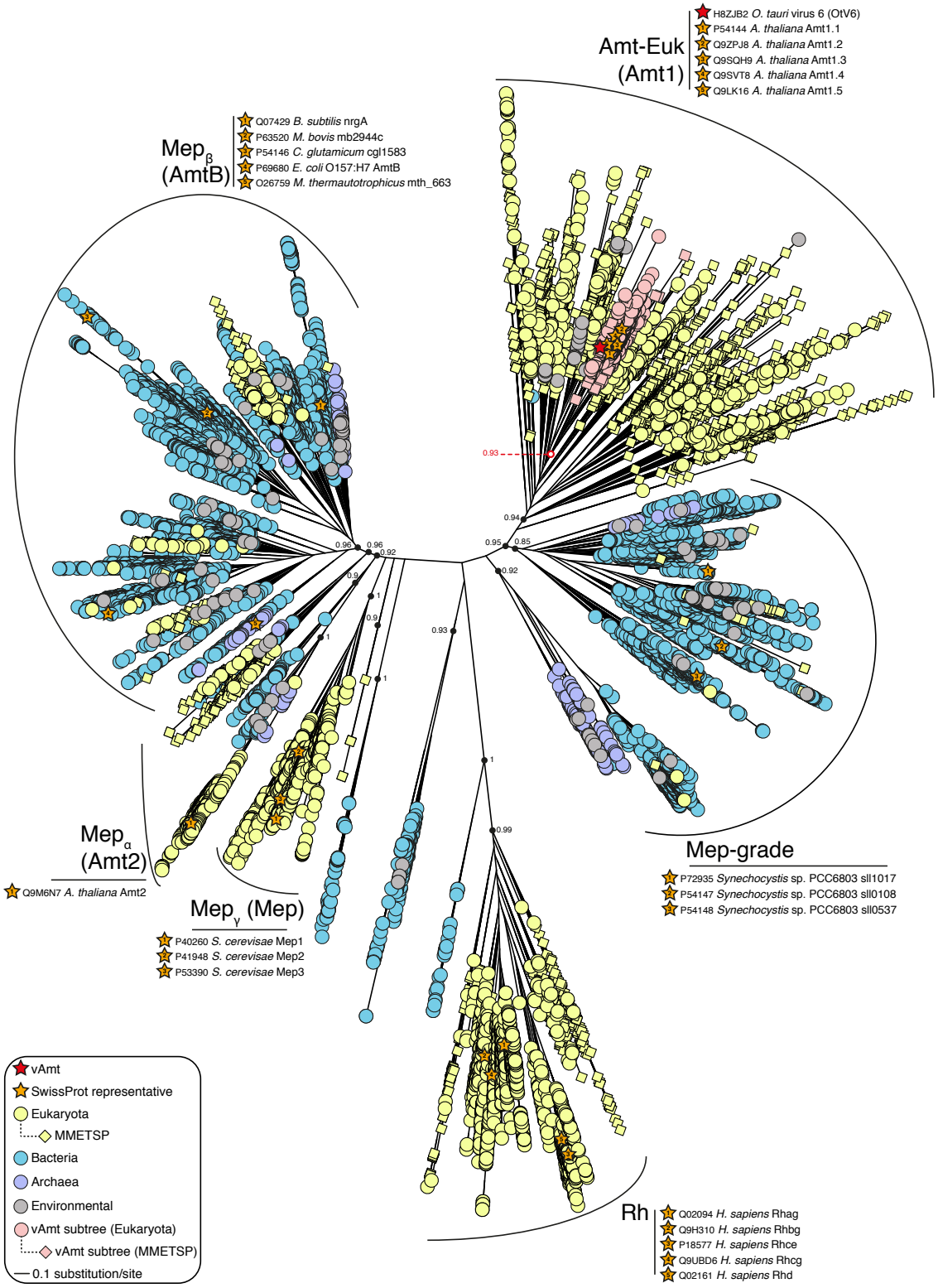

Fig. S11. Amt/Mep/Rh superfamily phylogenetic tree. This large-scale approximate maximum-likelihood phylogenetic tree was reconstructed under WAG+G model using FastTree v2.1 [13], based on a multiple alignment of 19,493 protein sequences totaling 374 sites. This phylogenetic tree is identical to the one displayed as Fig. 6A. Local support values were based on the Shimodaira-Hasegawa test [14]; support value of the vAmt subtree is indicated in red. The different Amt/Mep/Rh clades are outlined as in McDonald *et al.* [15], with the Mep clade encompassing the Mep$_\alpha$, Mep$_\beta$ and Mep$_\gamma$ groups. Each sequence is represented by a circle, or with a diamond in the case of sequences from the MMETSP project [16], and the color-code corresponds to the taxonomic information: purple, blue and yellow correspond to archaeal, bacterial and eukaryotic sequences, respectively. Grey circles correspond to environmental sequences with no available taxonomic information. The red star shows the phylogenetic position of the vAmt within the superfamily tree, and sequences grouping with the vAmt are highlighted in pink (all of which are eukaryotic), and were used to compute the maximum-likelihood phylogeny displayed as Fig. 6B and Fig. S11 (vAmt subtree). Orange stars represent SwissProt reviewed protein entries, and are listed with their corresponding sequence identifiers. The scale bar represents the number of estimated substitutions per site.

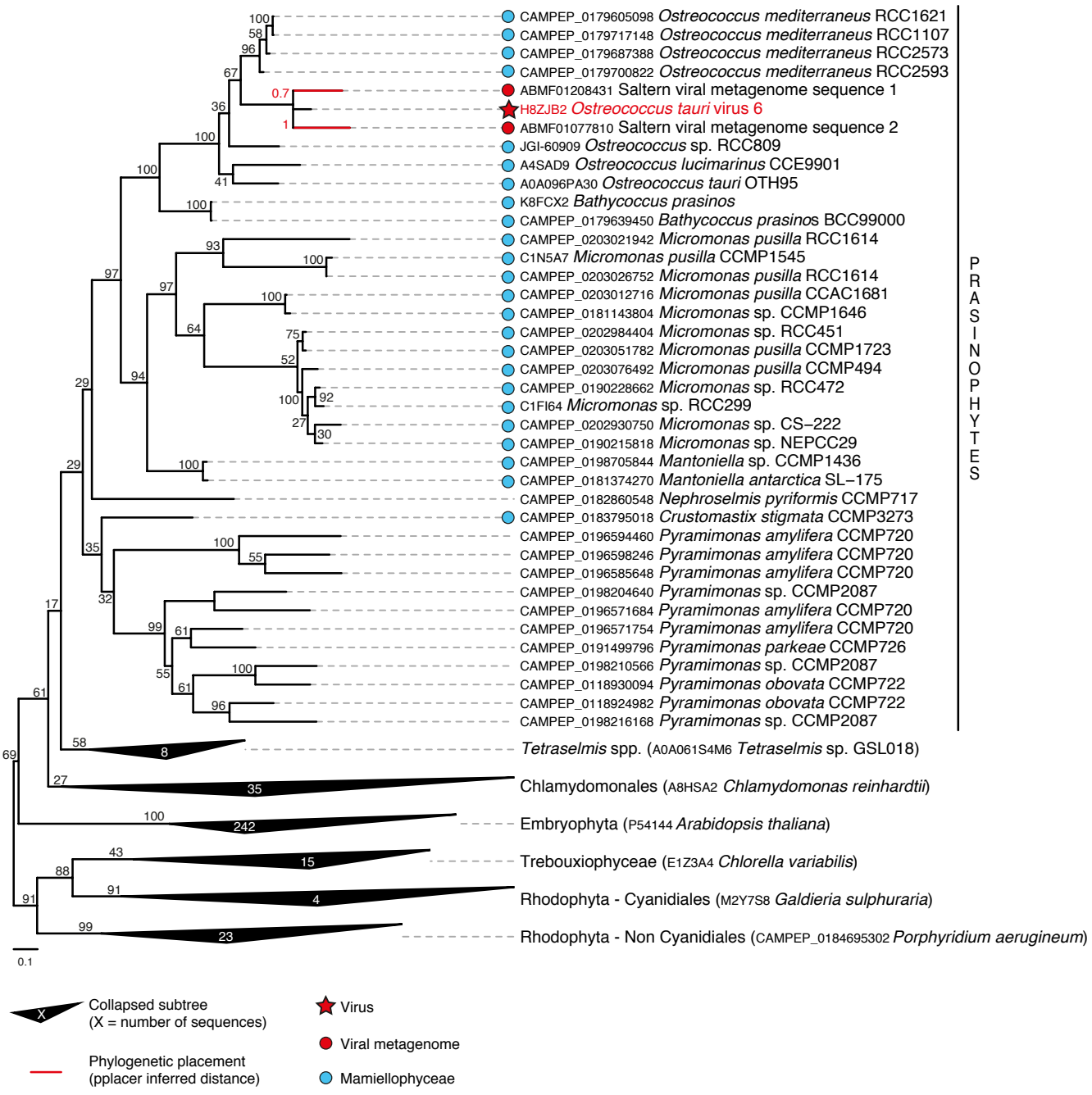

Fig. S12. vAmt evolutionary relationships with Amt-Euk homologs. This maximum-likelihood phylogenetic tree is identical to the one displayed as Fig. 5B but without collapsed branches within the prasinophytes, and was inferred under LG+I+G+F model using RAxML v8.2 [4], based on a multiple alignment comprised of 364 proteins totaling 429 sites. The vAmt is highlighted in red with a star, and branched within the prasinophyte lineage of green algae. Within the prasinophytes, Mamiellophyceae sequences are marked with a blue circle. Branches leading to Chlamydomonales, Embryophyta, Trebouxiophyceae and Rhodophyta were collapsed and the numbers indicated in parentheses are the total number of sequences present in the corresponding collapsed clade, listed with a representative sequence with corresponding UniProtKB identifier (or MMETSP [16] identifier in the case of non-Cyanidiales Rhodophyta). Code numbers in front of species names represent sequence identifiers from either the MMETSP transcriptomes (*O. mediterraneus*, *N. pyriformis* and *C. stigmata*), UniProtKB (*O. lucimarinus* and *O. tauri*) and the *Ostreococcus* sp. RCC809 genome project available at the DoE-Joint Genome Institute. Branch node supports were computed from 1000 non-parametric bootstrap replicates. The scale bar represents the number of estimated substitutions per site. The dashed branches show the phylogenetic placements of two short environmental sequences, which were placed onto the tree using pplacer v1.1 [17] with placement posterior probabilities indicated in red on their corresponding branch; both sequences originate from saltern viral metagenomes [18], and their GenBank sequence identifiers are provided (NCBI BioProject: PRJNA28353).

SI References

1. López-García P, Philippe H, Gail F, Moreira D (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc Natl Acad Sci USA* 100(2):697–702.

2. Bower S, et al. (2003) Preferential PCR amplification of parasitic protistan small subunit rDNA from metazoan tissues. *J Eukaryot Microbiol* 51(3):325–332.

3. Lanquar V, et al. (2009) Feedback inhibition of ammonium uptake by a phospho-dependent allosteric mechanism in *Arabidopsis*. *Plant Cell* 21(11):3610–3622.

4. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.

5. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30(4):772–780.

6. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25(15):1972–1973.

7. Rice P, Longden I, Bleasby A (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet* 16(6):276–277.

8. Krogh A, Larsson B, Von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305(3):567–580.

9. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948.

10. Holm L, et al. (2008) Searching protein structure databases with DaliLite v. 3. *Bioinformatics* 24(23):2780–2781.

11. Iraqui I, et al. (1999) Amino acid signaling in *Saccharomyces cerevisiae*: a permease-like sensor of external amino acids and F-Box protein Grr1p are required for transcriptional induction of the *AGP1* gene, which encodes a broad-specificity amino acid permease. *Mol Cell Biol* 19:989–1001.

12. Vaas LA, et al. (2013) opm: an R package for analysing OmniLog phenotype microarray data. *Bioinformatics* 29(14):1823–1824.

13. Price MN, Dehal PS, Arkin AP (2010) FastTree 2 —approximately maximum-likelihood trees for large alignments. *PloS One* 5(3):e9490.

14. Shimodaira H, Hasegawa M. (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16:1114–6.

15. McDonald TR, Dietrich FS, Lutzoni F (2012) Multiple horizontal gene transfers of ammonium transporters/ammonia permeases from prokaryotes to eukaryotes: toward a new functional and evolutionary classification. *Mol Biol Evol* 29(1):51–60.

16. Keeling PJ, et al. (2014) The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol* 12(6):e1001889.

17. Matsen FA, Kodner RB, Armbrust EV (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 11(1):538.

18. Dinsdale EA, et al. (2008) Functional metagenomic profiling of nine biomes. *Nature* 452(7187):629–632.