

SUPPLEMENTARY MATERIALS AND METHODS

Generation of transcriptomes from *Symbiodinium* type C1 populations

We queried transcriptome data of a thermo-sensitive *Symbiodinium* type C1 population (SM, aims-aten-C1-WSY) and a thermo-tolerant *Symbiodinium* type C1 population (MI, aims-aten-C1-MI) exposed to 27°C and 32°C in culture. Culture conditions, genotyping, experimental design, RNA extraction, transcriptome assembly, and transcriptome annotation are detailed by Levin *et al.* (2016). Briefly, samples were collected for transcriptome sequencing on day -1 (pre-heating; all samples acclimated to 27°C, n = 8 per population). On day 0, four samples per population were ramped at 0.5°C/hour to 32°C for the heat stress temperature treatment, while four samples per population remained at 27°C for the ambient temperature treatment. Samples (n = 4 per population per temperature treatment) were collected for transcriptome sequencing on days 9 and 13. The sampling time points for transcriptomics were determined by physiological measurements that detected thermal damage (reduced photosynthetic efficiency and increased ROS leakage from cells) to only the thermo-sensitive SM population (Levin *et al.*, 2016). Extracted RNA from each sampling time point was poly(A)+ purified, sequenced with an Illumina HiSeq2500 (100-bp single-end, strand-specific reads; $\sim 1 \times 10^7$ reads per sample), and assembled with Trinity (Grabherr *et al.*, 2011; Haas *et al.*, 2013) into *de novo* transcriptomes for each population. Transcripts in each *de novo* transcriptome that shared 99% sequence identity were collapsed into the longest representative transcript using cd-hit-est (Huang *et al.*, 2010) to remove redundancy. Transcriptomes were annotated with BLAST+ against SwissProt (e-value $\leq 10^{-5}$), UniRef90/TrEMBL (e-value $\leq 10^{-5}$), and Pfam-A databases (HMMER, domain noise cut-off). In the absence of a BLAST+ hit in SwissProt, annotation from UniRef90/TrEMBL was used, followed by Pfam-A.

Identification of viral and antiviral transcripts

A BLASTx bit score approach adapted from Boschetti *et al.* (2012) was performed to measure “foreignness” of transcripts for determining viral origin. Assembled transcripts from the *de novo* transcriptomes of the thermo-sensitive SM

and thermo-tolerant MI populations that received top BLAST+ hits to genes from eukaryotic viruses were aligned with BLASTx against partitioned SwissProt (curated) databases for archaea, bacteria, eukaryotes, and viruses. For eukaryotic virus transcripts that only received BLAST+ hits in UniRef90/TrEMBL, partitioned UniRef90/TrEMBL (non-curated) databases were used instead. For each sequence, the difference in bit score between the top BLASTx hit from the viral database and the top BLASTx hit from any non-viral database (h) was calculated. Only sequences with $h \geq 30$ (Boschetti *et al.*, 2012) and a bit score ≥ 50 (Roux *et al.*, 2011; Weynberg *et al.*, 2014; Wood-Charlson *et al.*, 2015) were considered to be of viral origin. The threshold of 30 was set for h as it provides the optimal trade-off between sensitivity and specificity in the detection of foreign genes (Boschetti *et al.*, 2012). To account for *Symbiodinium* genes that may be missing from the SwissProt or UniRef90/TrEMBL databases as well as to remove viral transcripts that have been integrated into the *Symbiodinium* genome from horizontal gene transfer events (keep only those that are from ongoing viral infections), BLASTx was run against both published *Symbiodinium* genomes (types B1 and F1) (Lin *et al.*, 2015; Shoguchi *et al.*, 2013). Again, only viral sequences with $h \geq 30$ were retained.

GC content and tri-nucleotide sequence statistics

(<http://www.genomatix.de/cgi-bin/tools/tools.pl>, last accessed July 2016) for the subset of viral transcripts and subset of non-viral transcripts in each transcriptome were converted from raw counts to frequencies (% occurrence) for subset comparisons (Supplementary Dataset 1). The program TransDecoder (<http://transdecoder.github.io/>, last accessed July 2016) was used to extract ORFs from the subset of viral transcripts and subset of non-viral (host) transcripts in each transcriptome. The EMBOSS tool CUSP (Rice *et al.*, 2000) was used to analyze codon usage frequencies of the ORFs in each subset (Supplementary Dataset 2).

Divergences in tri-nucleotide frequency matrices and codon usage matrices between subsets were determined by calculating the correlation coefficient for each matrix comparison in MATLAB (MathWorks, Inc.) (Supplementary Figure 1a-b).

Symbiodinium antiviral transcripts were identified by searching the annotation results for transcripts encoding known antiviral genes found in corals (Dunlap *et al.*, 2013) and for transcripts assigned gene ontology categories for

defense response to virus, negative regulation of viral genome, and negative regulation of viral genome replication. Gene ontology categories were assigned based on SwissProt annotation. In the absence of SwissProt gene ontology assignment, Pfam-A gene ontology assignment was used.

Differential expression analysis

Differential expression analysis for all transcripts in each population was conducted with RSEM (Li & Dewey, 2011) and edgeR (Chen *et al.*, 2014) as instructed in the standard Trinity pipeline (<https://github.com/trinityrnaseq/trinityrnaseq/wiki/Trinity-Differential-Expression>, last accessed July 2016)(Grabherr *et al.*, 2011; Haas *et al.*, 2013). Only transcripts with at least 2-fold differential expression at a maximum FDR of 0.001 between temperature treatments on at least one sampling time point were considered. Transcripts were not collapsed into Trinity 'genes' (transcript clusters based on shared sequence content) as not all transcripts in the same cluster always received annotation despite one or more of the transcripts aligning to a known viral gene - likely due to the limited amount of marine virus sequence information in databases (Brum & Sullivan, 2015; Wood-Charlson *et al.*, 2015). Transcript annotation was required for the HGT index approach.

PCR amplification of the +ssRNAV MCP gene

Total RNA was extracted and purified by Levin *et al.* (2016) using the RNeasy Plant Minikit (Qiagen) with an added on-column DNase treatment (Qiagen) and stored at -80°C. RNA was reverse transcribed into complementary DNA (cDNA) with the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). To obtain purified genomic DNA (gDNA), *Symbiodinium* cells (~5 x 10⁶) from thermo-sensitive SM and thermo-tolerant MI populations were washed with sterile media and then processed with the PowerPlant® Pro DNA Isolation Kit (MO BIO), which includes RNase-treatment of gDNA. Using ITS2 PCR primers from Stat *et al.* (2011), the ITS2 region was successfully amplified from the gDNA of the thermo-sensitive SM and thermo-tolerant MI populations, verifying that the gDNA was high quality and amplifiable.

MCP primer set 1 was designed for a long region of the MCP gene (970 nt) based on conserved sites in TR74740|c13_g1_i1 and TR74740|c13_g1_i2 (1F: 5' GCA CTT ACA AGG AAG CAG CGA GC 3', 1R: 5' CAT GAG CTC CAG CCT CCA ATG 3'). MCP primer set 2 was designed for a short region of the MCP gene (128 nt) based on conserved sites in TR74740|c13_g1_i1, TR74740|c13_g1_i2, and TR97578|c0_g1_i1 (2F: 5' TAT CAC CCG AGG TCC GTT TGG TTC 3', 2R: 5' TGA CTC TTG TCG TCC GAA TGA CTG TG 3'). Phusion® High-Fidelity PCR Master Mix with HF Buffer (New England BioLabs) was used as directed for 35 rounds of PCR with a primer annealing temperature of 63°C to amplify the long and short regions of the MCP gene from gDNA and cDNA of the thermo-sensitive SM and thermo-tolerant MI populations. Both primer sets successfully amplified the MCP gene from only cDNA (Supplementary Figure 2a). PCR products were sequenced by the Australian Genome Research Facility and confirmed to be from the MCP gene.

SUPPLEMENTARY DISCUSSION

Viral transcript diversity in *Symbiodinium* transcriptomes

The first discovery of potential sequences from pathogenic viruses (including *Mimiviridae*, *Phycodnaviridae*, *Poxviridae*, *Iridoviridae*, *Baculoviridae*, and *Herpesviridae*) in corals indicated that a diverse group of viruses infect the coral holobiont (Wegley *et al.*, 2007). Transcriptionally up-regulated viral A-type inclusion genes in heat stressed *Montastraea faveolata* and transcriptionally up-regulated transposon, pol-like protein, and reverse transcriptase genes in bleached *Montastraea faveolata* (DeSalvo *et al.*, 2008) suggested that coral-associated viruses are responsive to heat stress. Over 900 transcripts in each of our *Symbiodinium* type C1 transcriptomes received a top BLAST+ hit from at least one database to a gene from a eukaryotic virus. Similarly to findings in corals (Correa *et al.*, 2016; Wegley *et al.*, 2007), the transcripts shared homology with genes from *Mimiviridae*, *Phycodnaviridae*, *Poxviridae*, *Iridoviridae*, *Marseilleviridae*, *Asfarviridae*, *Baculoviridae*, *Herpesviridae*, *Polyomaviridae*, *Retroviridae*, and *Alvernaviridae* (*Dinornavirus*). However, the majority of these transcripts were not sufficiently

divergent from non-viral sequences and *Symbiodinium* genomic sequences ($h \geq 30$) (Boschetti *et al.*, 2012) to confidently assign them as true viral transcripts.

Following the BLASTx bit score approach (Supplementary Materials and Methods), 306 and 238 transcripts were concluded to be from viruses infecting the thermo-sensitive SM and thermo-tolerant MI populations, respectively. Viral transcripts expressed in both transcriptomes shared BLASTx homology with *Mimiviridae*, *Phycodnaviridae*, and *Dinornavirus* genes, though dominantly (>80%) with FNIP repeats in *Mimiviridae* genes. The diverse transcripts containing FNIP repeats had unknown functions since leucine-rich repeats (e.g., FNIP repeats) are common protein recognition motifs found in a broad variety of proteins (Kobe & Kajava, 2001). None of the FNIP repeat-containing transcripts within each transcriptome were identical. Furthermore, we used USEARCH (Edgar, 2010) with a minimum identity cut-off of 80% to collapse similar sequences and found reductions of only 254 to 172 and 211 to 155 transcripts in the thermo-sensitive SM and thermo-tolerant MI transcriptomes, respectively.

Genes containing FNIP repeats are prevalent in the genomes of *Mimivirus* and its relative *Cafeteria roenbergensis* virus (Fischer *et al.*, 2010; Suhre, 2005), thus an abundance of viral FNIP repeat-containing transcripts in our transcriptomes is not surprising. Furthermore, the relatively high proportion of viral FNIP repeat-containing transcripts compared to other *Mimiviridae*-like transcripts may be overestimated as a result of the limited amount of marine virus sequence information in databases preventing annotation of all viral genes in the transcriptomes (Brum & Sullivan, 2015; Wood-Charlson *et al.*, 2015) and/or an inability to confirm viral origin of other *Mimiviridae*-like transcripts. Phylogenetic analysis strongly indicates that an extensive number of genes were acquired by *Mimivirus* from its hosts as well as from bacteria infecting the same hosts (Moreira & Brochier-Armanet, 2008). Consequently, many *Mimiviridae*-like transcripts in our transcriptomes may not be sufficiently divergent ($h \geq 30$) from non-viral sequences and therefore would not be retained for further analysis. Moreover, although all *Mimivirus* genes are polyadenylated (Byrne *et al.*, 2009), certain dsDNA marine viruses contain a mix of poly(A)+ and poly(A)- genes (Tsai *et al.*, 2004). Accordingly, the breadth of gene types that can be retained through poly(A)+ purification cannot

be determined without knowledge of the polyadenylation rate exhibited by the specific viruses discovered in our study.

Other *Mimiviridae*-like transcripts detected in both transcriptomes shared homology with ankyrin-repeat protein genes, bifunctional E2/E3 enzyme genes, putative serine/threonine-protein kinase/receptor genes, a band 7 family protein gene, and uncharacterized protein genes. *Mimiviridae*-like transcripts uniquely detected in the thermo-sensitive SM transcriptome shared homology with a deoxynucleotide monophosphate kinase gene, an uncharacterized glycosyltransferase gene, a resolvase gene, and a transposase gene. A *Mimiviridae*-like gene detected only in the thermo-tolerant MI transcriptome was a N-glycosidase gene. One transcript with similarity to a *Phycodnaviridae* transcription factor gene was expressed in the thermo-sensitive SM transcriptome, and two transcripts with similarity to *Phycodnaviridae* uncharacterized protein genes were expressed in the thermo-tolerant MI transcriptome. Also, one transcript expressed in the thermo-sensitive SM transcriptome had similarity to an *Iridoviridae* uncharacterized protein gene.

As discussed in the main text, two transcripts in the thermo-sensitive SM transcriptome and one transcript in the thermo-tolerant MI transcriptome shared BLASTx homology with the *Dinornavirus* MCP gene. However, phylogenetic analysis revealed that the *Dinornavirus*-like MCP genes in our transcriptomes of type C1 *Symbiodinium* from South Molle Island and Magnetic Island (Levin *et al.*, 2016), partial *Dinornavirus*-like MCP genes in the metagenome of *Acropora tenuis* from Orpheus Island (Weynberg *et al.*, 2014), and partial *Dinornavirus*-like MCP genes in the virome of *Montastraea cavernosa* from Key West (Correa *et al.*, 2013) are substantially divergent from the known *Dinornavirus* MCP genes (Nagasaki *et al.*, 2005) as well as from one another (Supplementary Table 1, Supplementary Figure 3). The *Dinornavirus*-like MCP genes from the separate studies may therefore be members of different, unknown *Alvernaviridae* genera, potentially with distinctive host and/or geographic specificities.

SUPPLEMENTARY TABLES

Table S1.

The known *Dinornavirus* MCP genes compared to *Dinornavirus*-like MCP genes found in 'omics datasets from corals and *Symbiodinium*.

sequence name	sequence description	sequence source	dinoflagellate host	location	complete gene?	ORF length	accession number	reference
HcRNAV34	<i>Dinornavirus</i> HcRNAV34 MCP gene	<i>Dinornavirus</i> HcRNAV34	<i>Heterocapsa circularisquama</i>	Western Japan	Yes	1080 nt	NCBI: AB218608	Tomaru <i>et al.</i> , 2004; Nagasaki <i>et al.</i> , 2005
HcRNAV109	<i>Dinornavirus</i> HcRNAV109 MCP gene	<i>Dinornavirus</i> HcRNAV109	<i>Heterocapsa circularisquama</i>	Western Japan	Yes	1080 nt	NCBI: AB218609	Tomaru <i>et al.</i> , 2004; Nagasaki <i>et al.</i> , 2005
TR74740 c13_g1_i1	Homology to <i>Dinornavirus</i> MCP gene	Thermo-sensitive SM <i>Symbiodinium</i> transcriptome	Type C1 <i>Symbiodinium</i>	South Molle Island, Great Barrier Reef, QLD, AU	Yes	1077 nt	GenBank: KX538960, NCBI GEO: GSE77911	Levin <i>et al.</i> , 2016
TR74740 c13_g1_i2	Homology to <i>Dinornavirus</i> MCP gene	Thermo-sensitive SM <i>Symbiodinium</i> transcriptome	Type C1 <i>Symbiodinium</i>	South Molle Island, Great Barrier Reef, QLD, AU	Yes	1077 nt	GenBank: KX787934, NCBI GEO: GSE77911	Levin <i>et al.</i> , 2016
TR97578 c0_g1_i1	Homology to <i>Dinornavirus</i> MCP gene	Thermo-tolerant MI <i>Symbiodinium</i> transcriptome	Type C1 <i>Symbiodinium</i>	Magnetic Island, Great Barrier Reef, QLD, AU	No	471 nt	NCBI GEO: GSE77911	Levin <i>et al.</i> , 2016
729699.2	Homology to <i>Dinornavirus</i> MCP gene	<i>Acropora tenuis</i> metagenome	Hypothesized: Type C3 <i>Symbiodinium</i>	Orpheus Island, Great Barrier Reef, QLD, AU	No	297 nt	NCBI SRA run: SRR1207979	Weynberg <i>et al.</i> , 2014
527285.2	Homology to <i>Dinornavirus</i> MCP gene	<i>Acropora tenuis</i> metagenome	Hypothesized: Type C3 <i>Symbiodinium</i>	Orpheus Island, Great Barrier Reef, QLD, AU	No	279 nt	NCBI SRA run: SRR1210582	Weynberg <i>et al.</i> , 2014
503847.2	Homology to <i>Dinornavirus</i> MCP gene	<i>Acropora tenuis</i> metagenome	Hypothesized: Type C3 <i>Symbiodinium</i>	Orpheus Island, Great Barrier Reef, QLD, AU	No	258 nt	NCBI SRA run: SRR1210582	Weynberg <i>et al.</i> , 2014
GAIR4WK03FYXY4	Homology to <i>Dinornavirus</i> MCP gene	Heat-stressed <i>Montastraea cavernosa</i> virome	Hypothesized: Clade C or D <i>Symbiodinium</i>	Key West, Florida, USA	No	375 nt	NCBI SRA run: SRR493108	Correa <i>et al.</i> , 2013
GAIR4WK03GFJLN	Homology to <i>Dinornavirus</i> MCP gene	Heat-stressed <i>Montastraea cavernosa</i> virome	Hypothesized: Clade C or D <i>Symbiodinium</i>	Key West, Florida, USA	No	345 nt	NCBI SRA run: SRR493108	Correa <i>et al.</i> , 2013
GAIR4WK03F1XL6	Homology to <i>Dinornavirus</i> MCP gene	Heat-stressed <i>Montastraea cavernosa</i> virome	Hypothesized: Clade C or D <i>Symbiodinium</i>	Key West, Florida, USA	No	339 nt	NCBI SRA run: SRR493108	Correa <i>et al.</i> , 2013

Table S2.

Up-regulated (+) and down-regulated (-) viral transcripts in the thermo-sensitive SM transcriptome after 9 days at 32°C (fold ≥ 2 relative to 27°C on at least one sampling time point, FDR ≤ 0.001).

transcript	log ₂ (fold)	FDR	CPM	annotation	virus	bit score	<i>h</i> (against partitioned databases)	<i>h</i> (against <i>Symbiodinium</i> genomes)
TR74718 c0_g2_i1	+3.29	4.55E-11	1.93	Putative resolvase R771	<i>Mimiviridae</i>	129.80	49.69	129.80
TR74718 c0_g1_i1	+2.87	7.48E-16	3.89	Putative transposase R104	<i>Mimiviridae</i>	113.62	55.47	113.62
TR38522 c0_g3_i1	+1.86	3.47E-26	13.34	Uncharacterized protein L728	<i>Mimiviridae</i>	228.78	228.78	116.78
TR38522 c0_g1_i1	+1.77	1.83E-20	9.68	Uncharacterized protein L728	<i>Mimiviridae</i>	228.01	228.01	116.01
TR72194 c2_g1_i2	+1.72	4.85E-05	3.09	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	670.06	356.21	532.06
TR73034 c0_g1_i2	+1.36	7.34E-10	8.10	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	726.02	342.03	262.02
TR73034 c0_g1_i1	+1.33	8.45E-06	5.20	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	414.42	153.30	43.42
TR53463 c0_g1_i1	-0.82	2.25E-08	38.18	Uncharacterized protein R871	<i>Mimiviridae</i>	222.96	222.96	72.96
TR74740 c13_g1_i1	-2.84	3.01E-14	16378.33	Major viral capsid protein	<i>Dinornavirus</i>	179.10	179.10	179.10
TR74740 c13_g1_i2	-2.87	9.75E-15	2305.35	Major viral capsid protein	<i>Dinornavirus</i>	191.04	191.04	191.04

Table S3.

Up-regulated (+) and down-regulated (-) viral transcripts in the thermo-sensitive SM transcriptome after 13 days at 32°C (fold ≥ 2 relative to 27°C on at least one sampling time point, FDR ≤ 0.001).

transcript	log ₂ (fold)	FDR	CPM	annotation	virus	bit score	<i>h</i> (against partitioned databases)	<i>h</i> (against <i>Symbiodinium</i> genomes)
TR74370 c1_g3_i1	+5.77	6.61E-06	0.66	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	498.66	338.86	341.66
TR74718 c0_g2_i1	+3.03	1.83E-08	1.69	Putative resolvase R771	<i>Mimiviridae</i>	129.80	49.69	129.80
TR74718 c0_g1_i1	+2.81	2.19E-11	3.39	Putative transposase R104	<i>Mimiviridae</i>	113.62	55.47	113.62
TR42922 c0_g1_i1	+2.43	6.34E-06	1.14	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	537.24	184.48	334.24
TR38522 c0_g1_i1	+2.05	1.78E-26	11.85	Uncharacterized protein L728	<i>Mimiviridae</i>	228.01	228.01	116.01
TR66569 c0_g1_i2	+1.93	1.82E-06	1.80	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	635.88	323.14	343.88
TR72673 c23_g1_i1	+1.76	1.07E-04	1.44	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	527.27	39.68	196.27
TR38522 c0_g3_i1	+1.74	4.50E-21	15.99	Uncharacterized protein L728	<i>Mimiviridae</i>	228.78	228.78	116.78
TR73034 c0_g1_i1	+1.65	7.68E-10	5.01	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	414.42	153.30	43.42
TR73589 c0_g1_i1	+1.64	2.91E-05	2.39	Putative ankyrin repeat protein R784	<i>Mimiviridae</i>	53.91	53.91	53.91
TR72194 c2_g1_i2	+1.58	3.03E-06	3.51	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	670.06	356.21	532.06
TR130476 c0_g1_i1	+1.57	3.51E-04	1.51	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	515.70	350.48	211.70
TR31628 c0_g2_i1	+1.49	1.21E-05	2.17	Putative FNIP repeat-containing protein R636	<i>Mimiviridae</i>	470.23	85.87	222.23

TR73034 c0_g1_i2	+1.41	1.29E-13	8.47	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	414.42	153.30	262.02
TR74062 c9_g1_i2	+1.11	3.27E-04	3.07	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	516.45	155.22	314.45
TR37111 c0_g2_i1	-0.98	1.18E-04	4.48	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	571.47	308.83	451.47
TR53463 c0_g1_i1	-1.05	2.40E-13	33.57	Uncharacterized protein R871	<i>Mimiviridae</i>	222.96	222.96	72.96
TR37111 c0_g1_i1	-1.05	4.65E-04	2.65	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	529.52	243.00	386.52
TR65978 c2_g4_i1	-1.14	2.48E-04	2.53	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	454.38	138.63	361.58
TR12211 c0_g1_i1	-1.14	1.10E-04	2.62	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	225.26	225.26	136.26
TR22857 c0_g1_i1	-1.91	1.75E-04	1.05	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	583.82	241.46	426.82
TR73567 c8_g16_i1	-2.37	1.17E-04	0.82	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	709.72	394.73	578.72
TR74740 c13_g1_i1	-3.05	4.57E-20	9895.89	Major viral capsid protein	<i>Dinornavirus</i>	179.10	179.10	179.10
TR74740 c13_g1_i2	-3.10	1.63E-19	1389.39	Major viral capsid protein	<i>Dinornavirus</i>	191.04	191.04	191.04

Table S4.

Up-regulated (+) and down-regulated (-) viral transcripts in the thermo-tolerant MI transcriptome after 9 days at 32°C (fold ≥ 2 relative to 27°C on at least one sampling time point, FDR ≤ 0.001).

transcript	$\log_2(\text{fold})$	FDR	CPM	annotation	virus	bit score	h (against partitioned databases)	h (against <i>Symbiodinium</i> genomes)
TR51820 c0_g4_i1	+2.53	1.94E-08	1.45	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	684.03	245.73	349.03
TR61729 c0_g1_i1	+2.03	2.93E-17	6.46	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	644.37	525.74	169.37
TR62954 c0_g4_i1	+1.78	1.03E-11	4.13	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	590.00	132.50	404.00
TR61729 c0_g1_i2	+1.69	5.35E-12	5.41	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	486.05	189.12	118.05
TR61287 c9_g2_i1	+1.51	2.66E-04	1.98	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	668.61	265.76	327.61
TR18026 c0_g1_i1	+1.49	9.47E-05	1.80	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	465.98	78.58	318.98
TR63607 c7_g16_i7	+1.32	3.60E-04	1.67	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	651.17	385.07	532.17
TR13909 c1_g1_i1	+1.21	2.49E-04	2.70	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	781.43	383.20	558.43
TR10186 c0_g1_i1	+1.21	5.27E-05	4.24	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	731.71	183.31	565.71
TR62747 c9_g3_i2	+1.02	1.26E-05	4.72	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	607.00	298.13	246.00
TR61287 c8_g2_i1	+0.96	1.43E-04	6.22	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	457.93	90.12	174.93
TR63204 c5_g6_i1	+0.88	6.26E-04	6.83	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	612.01	273.86	291.01
TR46628 c0_g1_i1	-2.02	2.54E-07	2.83	Uncharacterized protein R883	<i>Mimiviridae</i>	80.88	80.88	83.17
TR105785 c0_g1_i1	-3.00	3.20E-05	0.79	Uncharacterized protein R617	<i>Mimiviridae</i>	127.10	71.26	42.40

Table S5.

Up-regulated (+) and down-regulated (-) viral transcripts in the thermo-tolerant MI transcriptome after 13 days at 32°C (fold ≥ 2 relative to 27°C on at least one sampling time point, FDR ≤ 0.001).

transcript	log ₂ (fold)	FDR	CPM	annotation	virus	bit score	<i>h</i> (against partitioned databases)	<i>h</i> (against <i>Symbiodinium</i> genomes)
TR51820 c0_g4_i1	+2.21	1.38E-04	1.05	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	684.03	245.73	349.03
TR79579 c0_g1_i1	+2.20	3.79E-04	0.96	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	366.65	69.70	58.65
TR61729 c0_g1_i2	+1.91	1.25E-13	4.61	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	486.05	189.12	118.05
TR61729 c0_g1_i1	+1.91	1.62E-21	7.44	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	644.37	525.74	169.37
TR10186 c0_g1_i1	+1.84	1.20E-12	4.30	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	731.71	183.31	565.71
TR63722 c16_g3_i4	+1.69	5.05E-04	1.24	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	591.87	214.10	479.87
TR17872 c0_g1_i1	+1.64	6.01E-09	3.37	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	617.32	335.82	478.32
TR62954 c0_g4_i1	+1.57	1.50E-07	3.69	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	590.00	132.50	404.00
TR65676 c0_g1_i1	+1.53	3.12E-06	3.19	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	390.52	90.51	127.52
TR63623 c3_g3_i8	+1.35	8.74E-04	1.79	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	594.59	191.01	478.59
TR61287 c8_g2_i1	+1.19	3.40E-08	6.53	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	457.93	90.12	174.93
TR62747 c9_g3_i2	+1.14	4.51E-06	4.46	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	607.00	298.13	246.00
TR27291 c2_g1_i3	+1.08	1.23E-04	4.18	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	773.28	467.90	637.28
TR63204 c5_g6_i1	+1.07	1.19E-04	6.29	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	612.01	273.86	291.01
TR60935 c0_g1_i1	-1.63	6.21E-10	4.48	Putative F-box and FNIP repeat-containing protein L60	<i>Mimiviridae</i>	729.10	345.11	408.10
TR46628 c0_g1_i1	-3.01	7.16E-07	1.39	Uncharacterized protein R883	<i>Mimiviridae</i>	83.17	83.17	83.17

Table S6.

Up-regulated (+) and down-regulated (-) antiviral transcripts in the thermo-sensitive SM transcriptome after 9 days at 32°C (fold ≥ 2 relative to 27°C on at least one sampling time point, FDR ≤ 0.001).

transcript	log ₂ (fold)	FDR	CPM	annotation
TR5690 c0_g1_i1	+1.43	2.26E-04	2.15	Baculoviral IAP repeat-containing protein 3
TR128451 c0_g1_i1	+1.36	1.15E-08	9.07	Influenza virus NS1A-binding protein
TR5690 c0_g1_i2	+1.28	1.71E-04	2.50	Baculoviral IAP repeat-containing protein 3
TR73376 c0_g1_i1	+1.24	5.58E-04	2.79	Influenza virus NS1A-binding protein
TR122761 c0_g1_i1	+1.08	2.35E-08	8.83	Influenza virus NS1A-binding protein homolog A
TR76594 c0_g1_i1	+0.93	7.68E-10	16.94	5'-3' exoribonuclease 2 homolog
TR61542 c0_g1_i1	+0.80	1.40E-04	9.24	Interferon-induced helicase C domain-containing protein 1

Table S7.

Up-regulated (+) and down-regulated (-) antiviral transcripts in the thermo-sensitive SM transcriptome after 13 days at 32°C (fold ≥ 2 relative to 27°C on at least one sampling time point, FDR ≤ 0.001).

transcript	log ₂ (fold)	FDR	CPM	annotation
TR5690 c0_g1_i1	+1.77	9.56E-08	2.83	Baculoviral IAP repeat-containing protein 3
TR5690 c0_g1_i2	+1.62	2.28E-08	3.56	Baculoviral IAP repeat-containing protein 3
TR128451 c0_g1_i1	+1.48	1.19E-08	11.61	Influenza virus NS1A-binding protein
TR122761 c0_g1_i1	+1.45	2.19E-07	10.68	Influenza virus NS1A-binding protein homolog A
TR61542 c0_g1_i1	+1.42	6.72E-11	9.51	Interferon-induced helicase C domain-containing protein 1
TR76594 c0_g1_i1	+1.19	8.01E-15	20.88	5'-3' exoribonuclease 2 homolog
TR73376 c0_g1_i1	+1.07	4.98E-04	3.16	Influenza virus NS1A-binding protein
TR24409 c0_g1_i1	+1.07	3.29E-06	5.15	Antiviral helicase SKI2
TR61681 c0_g1_i1	-2.46	3.67E-04	0.68	Baculoviral IAP repeat-containing protein 6

Table S8.

Up-regulated (+) and down-regulated (-) antiviral transcripts in the thermo-tolerant MI transcriptome after 9 days at 32°C (fold ≥ 2 relative to 27°C on at least one sampling time point, FDR ≤ 0.001).

transcript	log ₂ (fold)	FDR	CPM	annotation
TR61737 c0_g1_i1	-0.58	1.06E-03	8.29	Influenza virus NS1A-binding protein
TR68891 c0_g1_i1	-3.24	9.20E-10	2.98	Baculoviral IAP repeat-containing protein 7

Table S9.

Up-regulated (+) and down-regulated (-) antiviral transcripts in the thermo-tolerant MI transcriptome after 13 days at 32°C (fold ≥ 2 relative to 27°C on at least one sampling time point, FDR ≤ 0.001).

transcript	log ₂ (fold)	FDR	CPM	annotation
TR59840 c0_g2_i1	+2.82	2.07E-06	1.04	Baculoviral IAP repeat-containing protein 8
TR61737 c0_g1_i1	-1.08	7.43E-07	8.59	Influenza virus NS1A-binding protein
TR45815 c0_g1_i1	-1.40	1.42E-04	2.40	Ski2p
TR68891 c0_g1_i1	-3.16	1.45E-04	1.05	Baculoviral IAP repeat-containing protein 7

SUPPLEMENTARY FIGURES

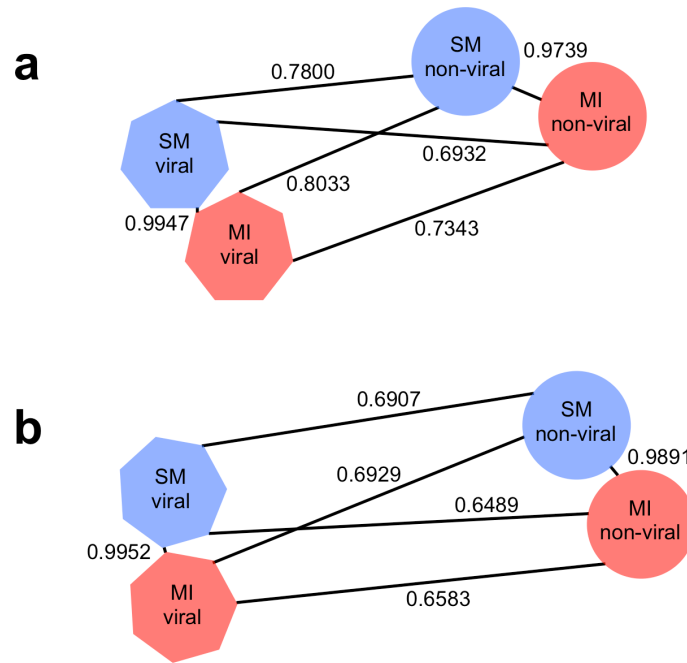


Figure S1.

Sequence characteristics diverge between virus and host transcripts.

The length of the line between the subsets of viral and non-viral (host) transcripts in the thermo-sensitive SM and thermo-tolerant MI transcriptomes represents the extent of divergence in **(a)** tri-nucleotide frequency and **(b)** codon usage. Correlation coefficients are reported for each comparison.

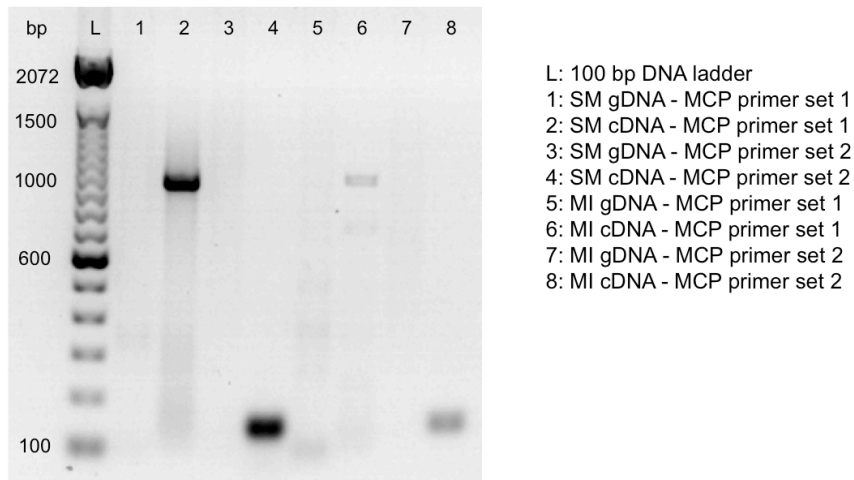


Figure S2.

The novel +ssRNAV MCP gene.

MCP PCR products were successfully amplified from cDNA, but not gDNA, of the thermo-sensitive SM and thermo-tolerant MI populations with MCP primer sets 1 and 2, supporting that transcripts TR74740|c13_g1_i1, TR74740|c13_g1_i2, and TR97578|c0_g1_i1 are of RNA virus and not host origin. The larger region of the MCP gene was amplified from cDNA of the thermo-tolerant MI population, confirming that the short length of TR97578|c0_g1_i1 (475 nt) was a result of incomplete *de novo* assembly due to low MCP transcript expression in the thermo-tolerant MI transcriptome.

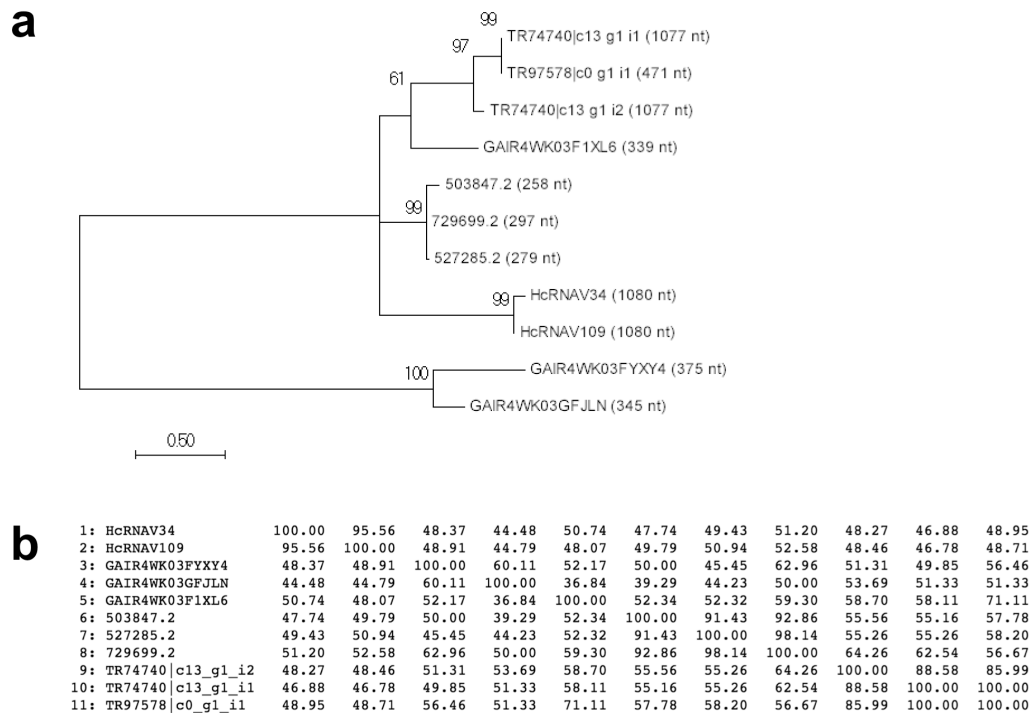


Figure S3.

Molecular phylogenetic analysis of +ssRNAV MCP genes.

Relationships between in-frame nt sequences of the *Dinornavirus* MCP genes and *Dinornavirus*-like MCP genes detailed in Supplementary Table 1 were assessed. (a) Evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (Tamura & Nei, 1993). The tree with the highest log likelihood (-2158.3565) is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches (Felsenstein, 1985). Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 2.9108)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 1.8733% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Codon positions included were 1st+2nd+3rd+Noncoding. All positions with less than 80% site coverage were eliminated. That is, fewer than 20% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 258 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar *et al.*, 2016). (b) A percent identity matrix was generated based on nt sequence alignments produced using MUSCLE (Edgar, 2004; McWilliam *et al.*, 2013).

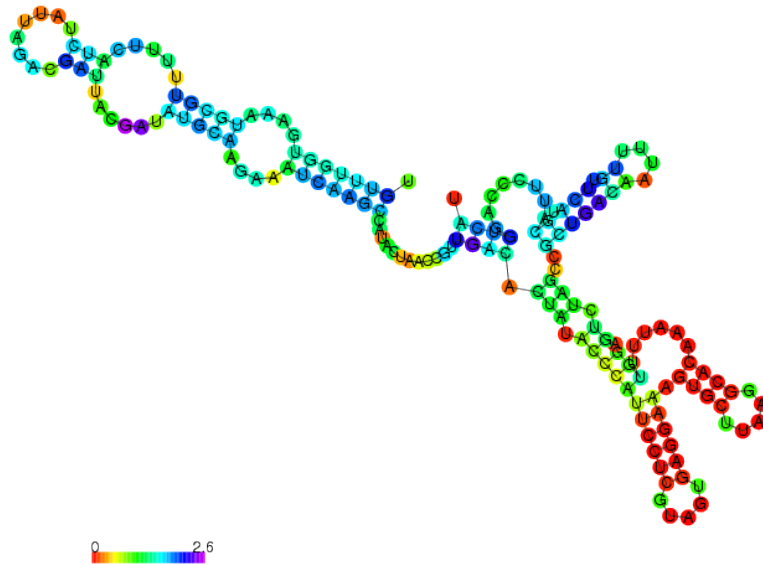


Figure S4.

The novel +ssRNAV IRES.

Secondary structure of the novel +ssRNAV IRES in TR74740|c13_g1_i1 and TR74740|c13_g1_i2 was predicted using VIPS and is similar to the cricket paralysis virus IRES (R score > 1.61, $P < 0.001$) (<http://140.135.61.250/vips/>, last accessed July 2016) (Hong *et al.*, 2013). The structure is colored by positional entropy.










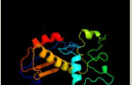
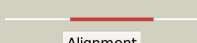


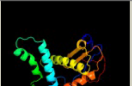

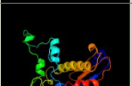




#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	d1u09a_	 Alignment		96.3	23	Fold: DNA/RNA polymerases Superfamily: DNA/RNA polymerases Family: RNA-dependent RNA-polymerase
2	c3n6mA_	 Alignment		95.0	26	PDB header: transferase Chain: A; PDB Molecule: rna-dependent rna polymerase; PDBTitle: crystal structure of ev71 rdrp in complex with gtp
3	c2b43D_	 Alignment		94.5	22	PDB header: viral protein Chain: D; PDB Molecule: non-structural polyprotein; PDBTitle: crystal structure of the norwalk virus rna dependent rna polymerase2 from strain hu/nlv/dresden174/1997/ge
4	c4nz0F_	 Alignment		94.2	18	PDB header: transferase Chain: F; PDB Molecule: genome polyprotein; PDBTitle: the emcv 3dpol structure at 2.8a resolution
5	c2jd1_	 Alignment		94.2	23	PDB header: hydrolase, transferase Chain: 1; PDB Molecule: picornain 3c, rna-directed rna polymerase; PDBTitle: crystal structure of the poliovirus precursor protein 3cd
6	c3nahC_	 Alignment		94.2	17	PDB header: transferase Chain: C; PDB Molecule: rna dependent rna polymerase; PDBTitle: crystal structures and functional analysis of murine norovirus rna-2 dependent rna polymerase
7	d1xr7a_	 Alignment		93.9	22	Fold: DNA/RNA polymerases Superfamily: DNA/RNA polymerases Family: RNA-dependent RNA-polymerase
8	d1xr5a_	 Alignment		93.5	21	Fold: DNA/RNA polymerases Superfamily: DNA/RNA polymerases Family: RNA-dependent RNA-polymerase
9	d1xr6a_	 Alignment		93.4	24	Fold: DNA/RNA polymerases Superfamily: DNA/RNA polymerases Family: RNA-dependent RNA-polymerase
10	c5i62A_	 Alignment		93.3	23	PDB header: viral protein, replication Chain: A; PDB Molecule: potential rna-dependent rna polymerase; PDBTitle: crystal structure of the insertion loop deletion mutant of the rna-2 dependent rna polymerase of a human picorbirnavirus

Figure S5.

The novel +ssRNAV RNA replicase.

The protein encoded by ORF-1 in TR74740|c13_g1_i1, which did not receive annotation from BLAST+, was modelled using Phyre2 in intensive modelling mode (<http://www.sbg.bio.ic.ac.uk/~phyre2>, last accessed July 2016) (Kelley *et al.*, 2015). The top ten results all matched RNA replicases (RNA-dependent RNA-polymerases) and, in all cases when specified, are from +ssRNAVs (results 2, 3, 4, 5, 6, 10).

SUPPLEMENTARY REFERENCES

- Boschetti C, Carr A, Crisp A, Eyres I, Wang-Koh Y, Lubzens E *et al.* (2012). Biochemical Diversification through Foreign Gene Expression in Bdelloid Rotifers. *PLoS Genet* **8**: e1003035.
- Brum JR, Sullivan MB. (2015). Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nature Rev Microbiol* **13**: 147-159.
- Byrne D, Grzela R, Lartigue A, Audic S, Chenivesse S, Encinas S *et al.* (2009). The polyadenylation site of *Mimivirus* transcripts obeys a stringent 'hairpin rule'. *Genome Res* **19**: 1233-1242.
- Chen Y, Lun ATL, Smyth GK. (2014). Differential expression analysis of complex RNA-seq experiments using edgeR. *Statistical analysis of next generation sequencing data*. Springer. pp 51-74.
- Correa AM, Ainsworth TD, Rosales SM, Thurber AR, Butler CR, Thurber RLV. (2016). Viral outbreak in corals associated with an in situ bleaching event: atypical herpes-like viruses and a new megavirus infecting *Symbiodinium*. *Front Microbiol* **7**.
- Correa AM, Welsh RM, Thurber RLV. (2013). Unique nucleocytoplasmic dsDNA and +ssRNA viruses are associated with the dinoflagellate endosymbionts of corals. *ISME J* **7**: 13-27.
- DeSalvo MK, Voolstra CR, Sunagawa S, Schwarz JA, Stillman JH, Coffroth MA *et al.* (2008). Differential gene expression during thermal stress and bleaching in the Caribbean coral *Montastraea faveolata*. *Mol Ecol* **17**: 3952-3971.
- Dunlap WC, Starcevic A, Baranasic D, Diminic J, Zucko J, Gacesa R *et al.* (2013). KEGG orthology-based annotation of the predicted proteome of *Acropora digitifera*: ZoophyteBase - an open access and searchable database of a coral genome. *BMC Genomics* **14**: 509-509.
- Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460-2461.
- Edgar RC. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-1797.
- Felsenstein J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783-791.
- Fischer MG, Allen MJ, Wilson WH, Suttle CA. (2010). Giant virus with a remarkable complement of genes infects marine zooplankton. *Proc Natl Acad Sci USA* **107**: 19508-19513.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I *et al.* (2011). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat Biotechnol* **29**: 644-652.

- Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J *et al.* (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* **8**: 1494-1512.
- Hong JJ, Wu TY, Chang TY, Chen CY. (2013). Viral IRES prediction system-a web server for prediction of the IRES secondary structure in silico. *PLoS ONE* **8**: e79288.
- Huang Y, Niu B, Gao Y, Fu L, Li W. (2010). CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**: 680-682.
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845-858.
- Kobe B, Kajava AV. (2001). The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* **11**: 725-732.
- Kumar S, Stecher G, Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* **33**: 1870-1874.
- Levin RA, Beltran VH, Hill R, Kjelleberg S, McDougald D, Steinberg PD *et al.* (2016). Sex, Scavengers, and Chaperones: Transcriptome Secrets of Divergent *Symbiodinium* Thermal Tolerances. *Mol Biol Evol* **33**: 2201-2215.
- Li B, Dewey CN. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.
- Lin S, Cheng S, Song B, Zhong X, Lin X, Li W *et al.* (2015). The *Symbiodinium kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis. *Science* **350**: 691-694.
- McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N *et al.* (2013). Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res* **41**: W597-W600.
- Moreira D, Brochier-Armanet C. (2008). Giant viruses, giant chimeras: the multiple evolutionary histories of *Mimivirus* genes. *BMC Evol Biol* **8**: 1.
- Nagasaki K, Shirai Y, Takao Y, Mizumoto H, Nishida K, Tomaru Y. (2005). Comparison of genome sequences of single-stranded RNA viruses infecting the bivalve-killing dinoflagellate *Heterocapsa circularisquama*. *Appl Environ Microbiol* **71**: 8888-8894.
- Rice P, Longden I, Bleasby A. (2000). EMBOSS: the European molecular biology open software suite. *Trends Genet* **16**: 276-277.
- Roux S, Faubladiere M, Mahul A, Paulhe N, Bernard A, Debroas D *et al.* (2011). Metavir: a web server dedicated to virome analysis. *Bioinformatics* **27**: 3074-3075.
- Shoguchi E, Shinzato C, Kawashima T, Gyoja F, Mungpakdee S, Koyanagi R *et al.* (2013). Draft assembly of the *Symbiodinium minutum* nuclear genome reveals dinoflagellate gene structure. *Curr Biol* **23**: 1399-1408.

Stat M, Bird CE, Pochon X, Chasqui L, Chauka LJ, Concepcion GT *et al.* (2011). Variation in *Symbiodinium* ITS2 sequence assemblages among coral colonies. *PloS ONE* **6**: e15854.

Suhre K. (2005). Gene and genome duplication in *Acanthamoeba polyphaga* Mimivirus. *J Virol* **79**: 14095-14101.

Tamura K, Nei M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* **10**: 512-526.

Tsai J-M, Wang H-C, Leu J-H, Hsiao H-H, Wang AH-J, Kou G-H *et al.* (2004). Genomic and proteomic analysis of thirty-nine structural proteins of shrimp white spot syndrome virus. *J Virol* **78**: 11360-11370.

Tomaru Y, Katanozaka N, Nishida K, Shirai Y, Tarutani K, Yamaguchi M *et al.* (2004). Isolation and characterization of two distinct types of HcRNAV, a single-stranded RNA virus infecting the bivalve-killing microalga *Heterocapsa circularisquama*. *Aquat Microb Ecol* **34**: 207-218.

Wegley L, Edwards R, Rodriguez-Brito B, Liu H, Rohwer F. (2007). Metagenomic analysis of the microbial community associated with the coral *Porites astreoides*. *Environ Microbiol* **9**: 2707-2719.

Weynberg KD, Wood-Charlson EM, Suttle C, van Oppen MJH. (2014). Generating viral metagenomes from the coral holobiont. *Front Microbiol* **5**: 2006.

Wood-Charlson EM, Weynberg KD, Suttle CA, Roux S, van Oppen MJH. (2015). Metagenomic characterization of viral communities in corals: mining biological signal from methodological noise. *Environ Microbiol* **17**: 10.