



HAL
open science

A Viral Immunity Chromosome in the Marine Picoeukaryote, *Ostreococcus tauri*

Sheree Yau, Claire Hemon, Evelyne Derelle, Hervé Moreau, Gwenaë L Piganeau, Nigel Grimsley

► **To cite this version:**

Sheree Yau, Claire Hemon, Evelyne Derelle, Hervé Moreau, Gwenaë L Piganeau, et al.. A Viral Immunity Chromosome in the Marine Picoeukaryote, *Ostreococcus tauri*. PLoS Pathogens, 2016, 12 (10), pp.e1005965. 10.1371/journal.ppat.1005965 . hal-01391068

HAL Id: hal-01391068

<https://hal.sorbonne-universite.fr/hal-01391068>

Submitted on 2 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH ARTICLE

A Viral Immunity Chromosome in the Marine Picoeukaryote, *Ostreococcus tauri*

Sheree Yau, Claire Hemon, Evelyne Derelle, Hervé Moreau, Gwenaël Piganeau, Nigel Grimsley*

Sorbonne Universités, UPMC Univ Paris 06, CNRS, Biologie Intégrative des Organismes Marins (BIOM, UMR 7232), Observatoire Océanologique, Banyuls sur Mer, France

* nigel.grimsley@obs-banyuls.fr



 OPEN ACCESS

Citation: Yau S, Hemon C, Derelle E, Moreau H, Piganeau G, Grimsley N (2016) A Viral Immunity Chromosome in the Marine Picoeukaryote, *Ostreococcus tauri*. PLoS Pathog 12(10): e1005965. doi:10.1371/journal.ppat.1005965

Editor: Shou-Wei Ding, University of California Riverside, UNITED STATES

Received: June 12, 2016

Accepted: September 29, 2016

Published: October 27, 2016

Copyright: © 2016 Yau et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Sequence data used in this study can be found in the GenBank data libraries. *O. tauri* RCC4221 chromosome sequences are under accession numbers CAID0100001.2 to CAID0100020.2 except for the updated chromosome 19 sequence (this study), which is available from <http://www.php.obs-banyuls.fr/publications/data/2/>. Transcriptome data is available from PRJNA344946. The updated genome sequence and annotation of OtV5 is under accession EU304328.2. Gene models and annotations for *O. tauri* RCC4221 are also available from the Online Resource for Community

Abstract

Micro-algae of the genus *Ostreococcus* and related species of the order Mamiellales are globally distributed in the photic zone of world's oceans where they contribute to fixation of atmospheric carbon and production of oxygen, besides providing a primary source of nutrition in the food web. Their tiny size, simple cells, ease of culture, compact genomes and susceptibility to the most abundant large DNA viruses in the sea render them attractive as models for integrative marine biology. In culture, spontaneous resistance to viruses occurs frequently. Here, we show that virus-producing resistant cell lines arise in many independent cell lines during lytic infections, but over two years, more and more of these lines stop producing viruses. We observed sweeping over-expression of all genes in more than half of chromosome 19 in resistant lines, and karyotypic analyses showed physical rearrangements of this chromosome. Chromosome 19 has an unusual genetic structure whose equivalent is found in all of the sequenced genomes in this ecologically important group of green algae.

Author Summary

We propose that chromosome 19 of *O. tauri* is specialized in defence against viral attack, a constant threat for all planktonic life, and that the most likely cause of resistance is the over-expression of numerous predicted glycosyltransferase genes. *O. tauri* thus provides an amenable model for molecular analysis of genome evolution under environmental stress and for investigating glycan-mediated host-virus interactions, such as those seen in herpes, influenza, HIV, PBCV and mimivirus.

Introduction

Eukaryotic micro-algae of the genus *Ostreococcus* and related species of the order Mamiellales are globally distributed in the photic zone of the world's oceans where they contribute to fixation of atmospheric carbon and production of oxygen, besides providing a primary source of nutrition in the food web [1–3]. Their tiny size (1–3 μm), simple cells (one chloroplast, one

Annotation of Eukaryotes (ORCAE) under *Ostreococcus tauri* V2.

Funding: This work was supported by the French Agence Nationale de la Recherche grants REVIREC (ANR-12-BSV7-0006-01, coordinator NG) and DECOVIR (ANR-12-BSV7-0009, coordinator Y. Desdevises). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

mitochondrion), ease of laboratory culture and extremely small genomes, several of which have been completely sequenced [4–8], render them attractive as models for marine ecology [9–11], cell biology [12] and evolution in the green lineage [4,13]. Typical features of the highly streamlined genomes of the Mamiellales include a higher GC content (48–64% [7,14], 59% for *O. tauri*) than higher plants (41%, [15]) and two unusual “outlier” lower GC chromosomes that can carry higher proportions of transposons and genes predicted to originate from prokaryotes (see [16] for a recent review). In the type species *O. tauri*, these two chromosomes are chromosome 2 (the big outlier chromosome or BOC) and chromosome 19 (the small outlier or SOC), both of which greatly vary in size between strains isolated from the environment [17].

In marine environments, unicellular organisms account for the largest biomass, but they are outnumbered by about ten to one by viruses, which can lyse host cells and thereby contribute to biogeochemical cycles by promoting the turnover of plankton populations [18–20]. While large DNA viruses infecting algae have been known for many years (see [21,22] for reviews), including viruses of *Micromonas pusilla*, a member of the Mamiellales [23], viruses of *Ostreococcus* spp. were discovered more recently. Complete genomes of many viruses infecting Mamiellales (genus *Prasinovirus*) are now available [24–27]. Given the size of an *Ostreococcus* cell (about 1 μm) prasinoviruses are huge icosahedral particles, about one eighth of the host cell diameter between faces (110 nm), carrying a repertoire of about 250 genes. Resistance to prasinoviruses arises spontaneously in cultured algal lines [28] and their complex host–strain specificity patterns [29,30] witness their involvement in the planktonic ‘arms race’ that assures for rapid turnover and evolution in micro-algal populations (see [31] for a review). Viruses of algae in general and prasinoviruses in particular play an important role in controlling phytoplankton populations [32–36], the effectiveness of host defence responses thus determines the fate of algal populations.

Regrowth of eukaryotic microalgae following viral lysis with large dsDNA viruses has been observed in phylogenetically diverse lineages [37,38] and such virus-resistant lines can be stable in culture [28,38,39]. In the diplont marine bloom-forming haptophyte *Emiliania huxleyii* viral infection of diploid cells led to the selection of haploid cells which were resistant to infection [40]. Cell lines of the toxic bloom-forming dinoflagellate *Heterocapsa circularisquama* became resistant to viral infection by HcRNAV after co-culture with this single-stranded RNA virus. In this case, resistance appeared to be reversible and a variable proportion of resistant cells in culture harboured viruses and produced particles [41]. Some of these authors suggest that changes in the host cell surface might occur, or that non-infectious defective particles in culture might occupy available receptor sites on host cells, but no specific molecular mechanism has been demonstrated.

In *O. tauri*, Thomas et al. [28] observed two kinds of virus-resistant cell lines, either lines which were resistant but chronically infected, producing viruses at a low level in culture (“resistant producers” referred to as R^P), or cell lines devoid of viruses and immune to re-infection (“resistant non-producers”, or R^{NP}). Here we aimed to unravel the molecular mechanisms underlying viral resistance using clonal lines of *O. tauri* RCC4221 [4] and its virus, OtV5. We produced numerous independent OtV5-resistant lines to examine host and viral gene expression in detail with RNA-Seq technology and to test whether altered gene expression could be responsible for the resistance phenotypes.

Results

Generation of virus-resistant *O. tauri*

The experimental schema in Fig 1 shows how independent clonal lines of *O. tauri* resistant to OtV5 were generated from a clonal starting population. The parent OtV5-susceptible *O. tauri*

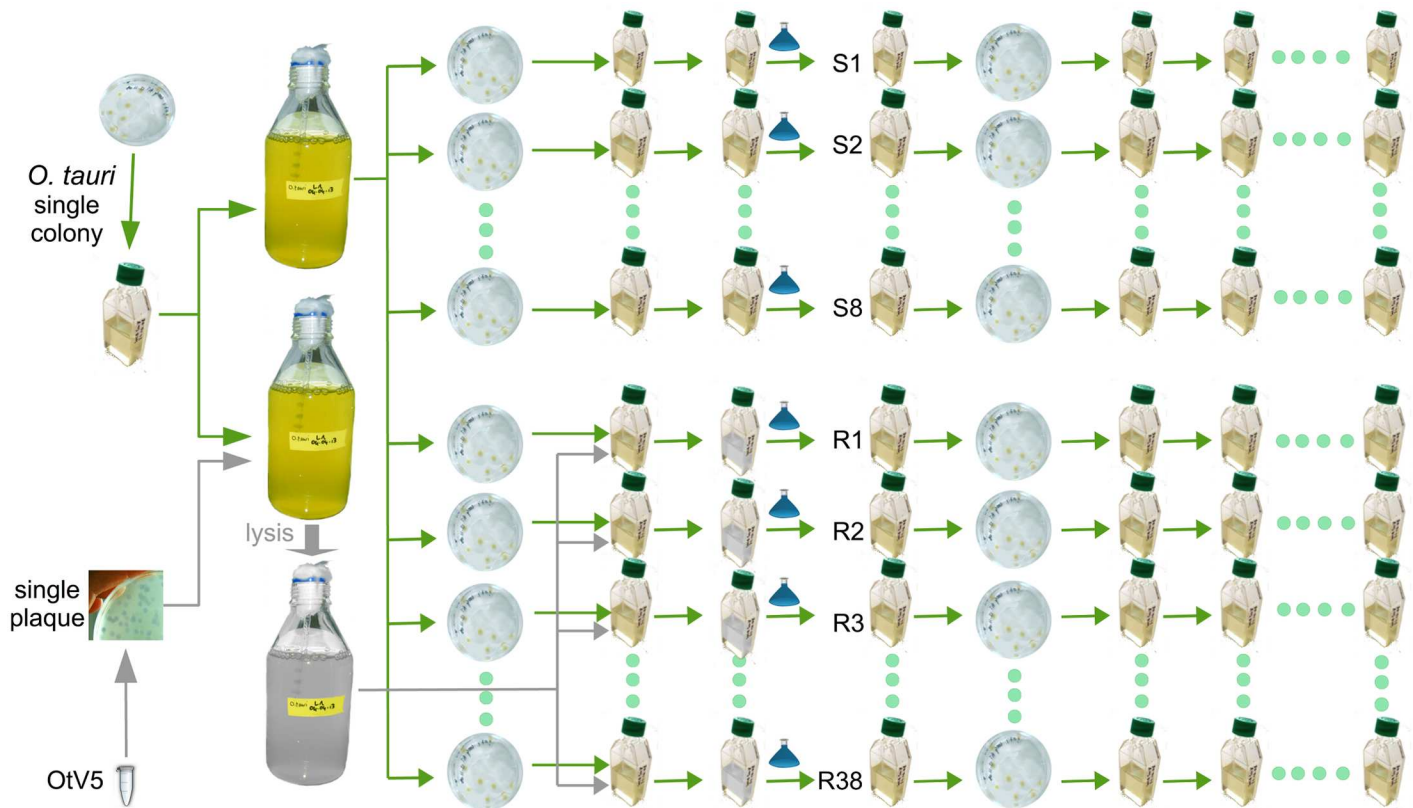


Fig 1. Experimental strategy for production of *Ostreococcus tauri* clonal lines susceptible or resistant to OtV5. A single colony of *O. tauri* was used to produce two 1-litre cultures of *O. tauri* cells, one used to prepare DNA for genome re-sequencing and to produce 46 independent clonal lines and another that was lysed by clonal OtV5. The viral lysate was subsequently used to inoculate 38 small flasks. Fresh medium (small blue flasks) was added to each lysate or control flask, and after ~1 week OtV5-resistant cells grew. A single colony from all lines was randomly chosen after plating and maintained in liquid culture for transcriptome sequencing and further analyses.

doi:10.1371/journal.ppat.1005965.g001

was subcultured into 46 parallel independent culture lines (one set of 30 and another set of 16). OtV5 was introduced into 38 of the *O. tauri* lines while eight were maintained as non-infected controls. All cultures with added OtV5 lysed, appearing clear of visible cells within 3–5 days. The cultures exposed to OtV5 showed regrowth of cells after approximately one week and, upon re-exposure to OtV5, were subsequently resistant to lysis. Periodic re-testing for OtV5 resistance showed viral resistance persisted over the almost two-year duration of the study while the control lines continued to be susceptible (8 susceptible control lines were maintained alongside the resistant lines shown in Fig 1, and at each of the 10 dates shown 8/8 lines remained susceptible to OtV5; at the end of the time course, all of the remaining 36 lines were found to be resistant). Resistant lines were periodically tested for production of infective OtV5 by taking cell-free culture medium from resistant lines, adding it to the susceptible parent and observing lysis. The majority of resistant lines were R^P at the start of the experiment, however, the proportion of R^P diminished over time so that R^{NP} lines were more abundant at the end of the study (Fig 2).

Differentially expressed genes are clustered on chromosome 19 and mainly related to carbohydrate metabolism

Twenty-three *O. tauri* transcriptomes were sequenced from the experiment including 19 virus-resistant and four susceptible. To test if OtV5-resistance was linked to differential gene

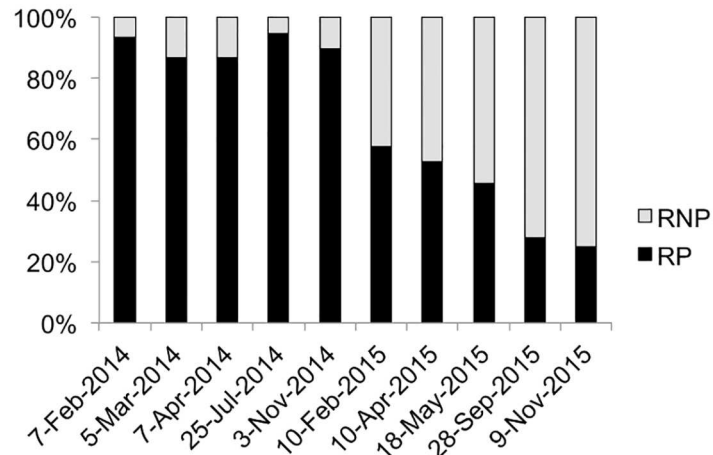


Fig 2. Proportion of resistant producer (RP) and resistant non-producer (RNP) lines over the course of the study. Percentage of RP and RNP in resistant lines on the ordinate and date when virus production test was performed on the abscissa.

doi:10.1371/journal.ppat.1005965.g002

expression and to identify the genes involved, we compared four resistant to four susceptible transcriptomes from matching RNA library batches (S1 Table). At this early point in the study all but two lines were R^P, so resistant lines used in the comparison were R^P, however over the course of the study several lines of evidence indicated R^P and R^{NP} were probably due to similar molecular mechanisms (discussed below). Over 95% (7432 of the 7749) of *O. tauri* chromosomal genes analysed were transcribed to some extent (≥ 10 aligned reads in at least one sample) in both resistant and susceptible *O. tauri* lines suggesting that the majority of genes were expressed at the time of sampling.

A total of 170 *O. tauri* chromosomal genes were significantly differentially transcribed in resistant lines of which 103 were up-regulated and 67 down-regulated (Fig 3, and see S2 and S4 Tables for full gene descriptions), which represents only 2% of all expressed *O. tauri* genes. Most strikingly, 49 differentially transcribed genes, representing almost a third of differentially transcribed genes were concentrated on chromosome 19, the SOC (Fig 3A). These genes on chromosome 19 also had overall larger log₂ fold change values indicating larger changes in expression levels on genes from this chromosome than the others (Fig 3B). Predicted glycosyltransferases (GTs) appear to be enriched on chromosome 19 compared to the other chromosomes, the majority of which were differentially transcribed in OtV5-resistant lines (Fig 3C). The only other chromosome encoding GTs regulated in virus resistant lines was chromosome 2, the BOC, where the GTs were down-regulated. Annotated genes on chromosome 19 are known to belong to few functional categories [5] relating to surface membrane proteins, the building of glycoconjugates, as well as methyltransferases (MTs) (Fig 4, S2 Table). The majority (16 of 22) of differentially transcribed carbohydrate transport and metabolism genes (Fig 3D) were located on chromosome 19. Furthermore, almost half of the differentially transcribed genes of unknown function (29 of 54) (Fig 3D) were also encoded by chromosome 19. Thus, genes involved in virus immunity in *O. tauri* are preferentially encoded by the SOC; these genes are strongly associated with carbohydrate modification and metabolism and their specific regulation was associated with viral resistance.

Other functional categories that were significantly differentially transcribed were related to translation, transcriptional regulation (chromatin remodelling, RNA modification, transcription factors), protein modification and turnover, amino acid transport and modification, and

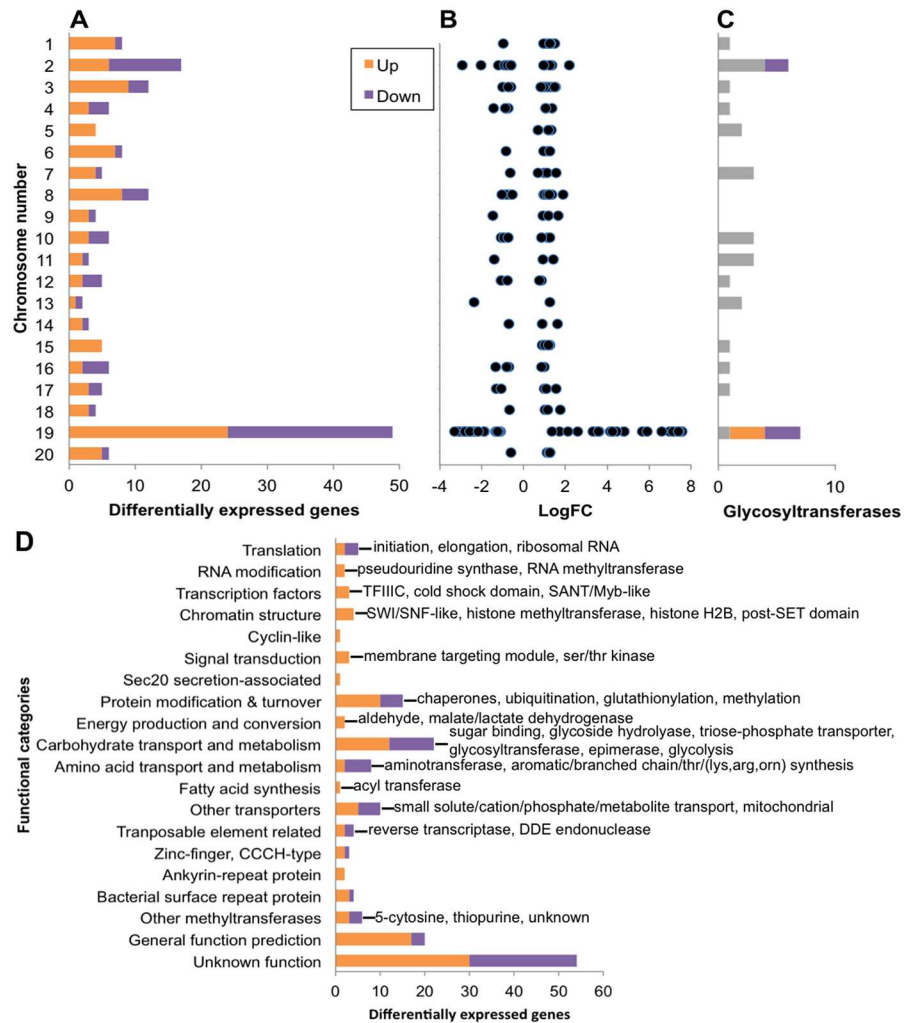


Fig 3. Differential gene expression in OtV5-resistant compared to susceptible control *O. tauri*. Up (orange) and Down (purple) regulated genes in virus resistant lines. **(A)** Counts of differentially transcribed genes per chromosome. **(B)** Plot of the log₂ fold change values of differentially transcribed genes per chromosome. **(C)** Predicted glycosyltransferase gene counts per chromosome based on matches to InterPro domains and families associated with glycosyltransferases. Genes with no difference in regulation are shown in grey. **(D)** Counts of differentially transcribed genes grouped according to functional categories shown on the ordinate. Labels describe the genes in each category (full gene descriptions in S2 and S4 Tables). Genes located on the large inverted duplicated region on chromosome 19 were only counted once.

doi:10.1371/journal.ppat.1005965.g003

other transporters (Fig 3C). Amino acid biosynthesis genes involved in four different pathways were down-regulated in resistant lines, implicating a decrease in a broad range of amino acids and their downstream metabolites during viral immunity (S4 Table). Ribosomal subunits were also under-expressed suggesting a slowing of growth-related processes. Genes involved in expression regulation were likely involved in maintenance of the virus-resistant state. In particular, histone modification genes were all over-transcribed in the resistant lines pointing to a key role for chromatin restructuring in resistance. Proteasome-mediated protein degradation (ubiquitination) as well as post-translational modification (glutathione-associated and chaperones) were up-regulated and may be part of the viral immunity response or translational-level regulation. Transporters whose substrates were related to undefined small, generally inorganic molecules had mixed regulation with a few cases of transporters with related functions being

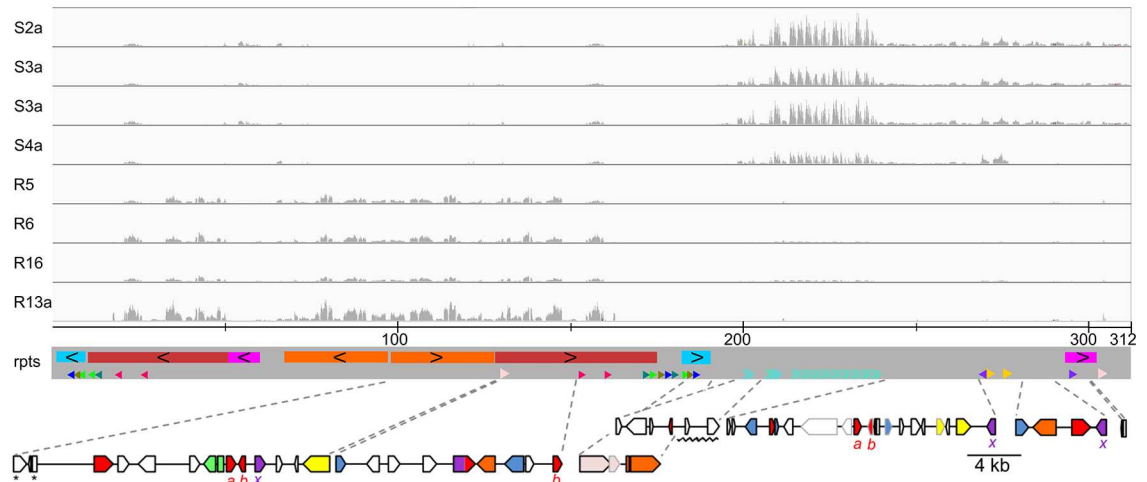


Fig 4. Bipartite structure of chromosome 19. Genes over-transcribed in virus-resistant lines were all located in a region of ~180 kb (on the left) containing long inverted repeats while under-transcribed genes were on a ~100 kb region (on the right). Tracks from top to bottom: alignment of RNA-Seq reads from susceptible (S2a–S5a) and resistant (R5, R6, R16 and R13a) samples (read depths have the same scale) onto chromosome 19 assembled from PacBio reads; scale in kb; (rpts) rectangles of the same colour on a grey background represent large duplicated blocks with black chevrons indicating inverted relative orientation and small arrowheads and chevrons of the same colour and style represent short repeats and their relative orientation; at the bottom, genomic maps zooming-in on blocks encoding genes under-transcribed (right) and over-transcribed (left) in virus-resistant lines (blue—methyltransferase, red—glycosyltransferase, green—epimerase/dehydratase, purple—sugar phosphate transporter, orange—other sugar metabolism, light pink—cell surface repeat lipoprotein, yellow—transposon-related, white—ORFAn or other functions). N.B. the coloured rectangles on the grey background are **not** related to those used in the gene maps beneath. Significantly differentially transcribed genes are outlined in black (see S2 Table for full annotations of these genes) and those not differentially regulated are outlined in light grey. The gene map region underlined by a zigzag is part of a series of tandem repeats shown by light blue chevrons (see S3 Fig). *a*, *b*, *x*. Genes with the same subscript letter are putative paralogues. *These genes are duplicated as they span the ends of the adjacent inverted blocks but the first gene (*osta19g00040*) includes a short tract of unique intervening sequence.

doi:10.1371/journal.ppat.1005965.g004

regulated in opposing directions. For example, a transmembrane phosphate transporter (*osta06g00210*) over-expressed while another calcium-dependent phosphate transporter (*osta17g00940*) was under expressed (S4 Table). Altered transcription of transporters may modulate the available substrate pool required by OtV5 curbing viral replication. However, these genes show lower significance levels of differential expression and may be involved in processes that are not directly relevant for resistance.

The bipartite structure and transcription profile of chromosome 19 suggests a regulatory switch

Re-sequencing of the *O. tauri* parent line with single molecule PacBio technology was able to resolve a large inverted repeat region (LIRR) on chromosome 19 totalling ~180 kb (Fig 4) that was predicted to exist from previous work [4]. The assembled PacBio contig was 311,428 bp, in agreement with its expected size from gel mobility (Fig 5). Notably, all over-transcribed genes in the resistant lines from the SOC were located in this region (Fig 4). Differentially transcribed genes on other chromosomes (S4 Table) were in some cases adjacent to each other, but were not grouped together as on chromosome 19. Remarkably, transcription of blocks of genes in the LIRR was almost completely suppressed in controls indicating close to half the SOC was silenced in the OtV5-susceptible wild-type state. By contrast, in resistant lines, these genes were highly transcribed, having among the highest mean fragment counts of the differentially

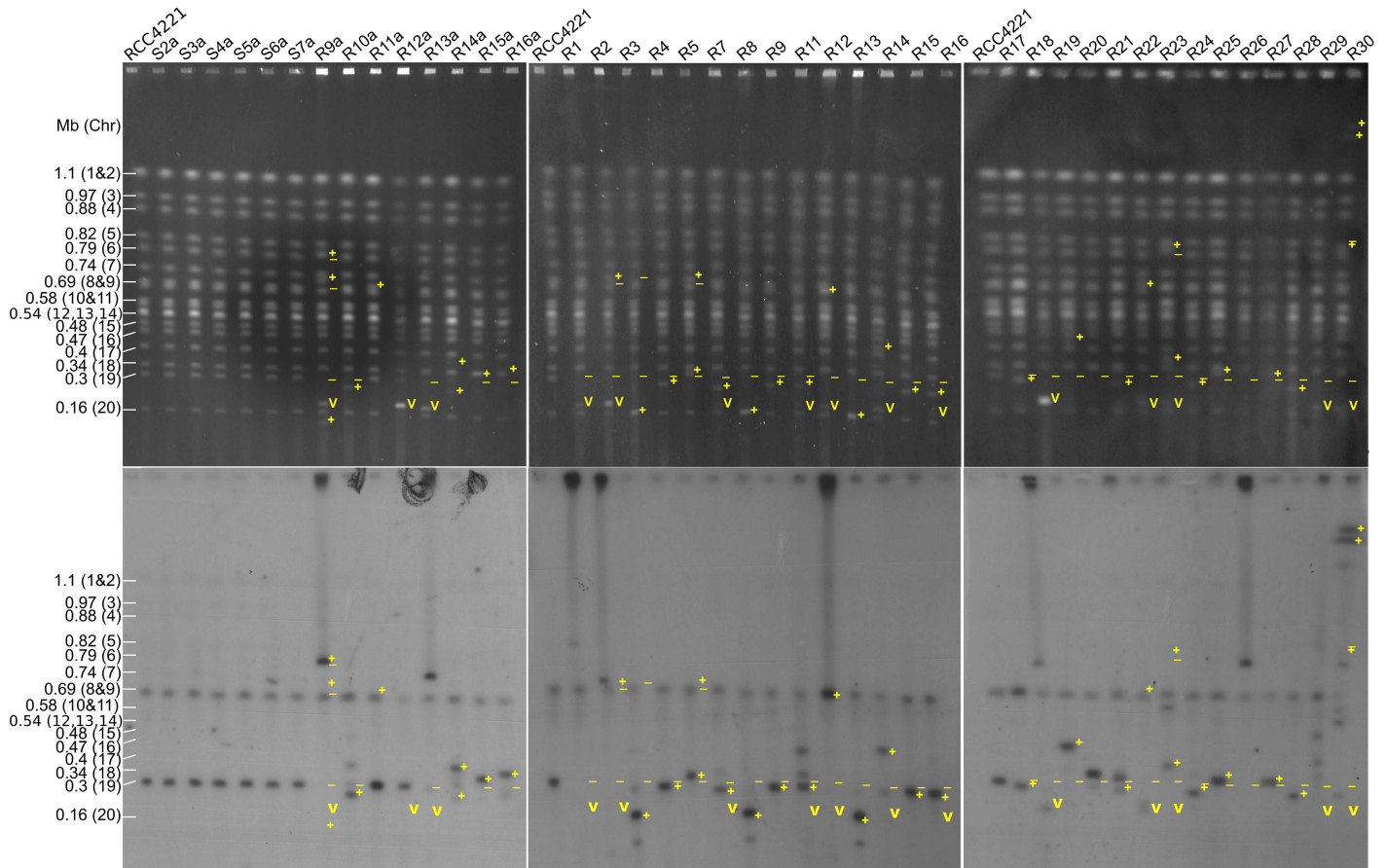


Fig 5. Pulsed-field gel electrophoresis (PFGE) analysis of susceptible and resistant lines ethidium bromide stained (top) and corresponding hybridization (bottom) using a probe for a gene on chromosome 19. Chromosome number and size are indicated on the left. Symbols on the upper gel images show, at positions observed by the ethidium-bromide fluorescence: - : absence of a band that was present in controls, +: presence of a band that was not present in controls, v: presence of a band at the expected size of the OtV5 genome. To aid the comparison, symbols from the upper images of ethidium-stained gels are shown at the same positions on the bottom Southern blot autoradiograph images. Note that since the probe hybridizes to a specific part of chromosome 19, occasionally the band of differing mobility does not correspond to the band identified by radioactive labelling, witnessing fragmentation of chromosome 19 after deletions, insertions, or translocations that may have occurred.

doi:10.1371/journal.ppat.1005965.g005

transcribed genes (S2 Table). The co-transcription of genes on the LIRR suggests they were under the control of the same regulatory factors.

Re-sequencing was also able to resolve a tandem repeat region (TRR) comprised of ~2,255 bp monomeric repeats sharing 97–99% identity (S1 Fig), which was under-transcribed in resistant lines but still showed some transcription with high sample–sample variance (Fig 4). Although small genes of unknown function (ostta19g00035 and ostta19g00240) occurred in the TRR, the sequences outside of these genes appeared transcribed, suggesting a functional role. The repeats were interrupted at two loci by unique sequences; (1) a 2,232 bp putative terminal repeat retrotransposon in miniature (TRIM) possessing long terminal repeats (LTRs) and (2) a 3,354 bp sequence of unclear origin (S1 Fig). The putative TRIM is likely derived from the complete 5,537 bp LTR retrotransposon, *retrostreo2* on chromosome 8, with which it shares 99% identity in the LTRs. Extensive deletions in the intervening gene sequences suggests the putative TRIM is non-autonomous and is an example of LTR-transposon miniaturisation in *O. tauri*, as well as transposition between chromosomes. Tandem repeats interrupted by transposons are hallmarks of centromeric repeats [42,43]. However, the repeat sequence was

not found on other chromosomes, as would be expected in centromere recognition sites, leaving the functional significance of the TRR unclear.

We conducted an exploratory analysis of all sequenced transcriptomes (S1 Table) to see if the chromosome 19 genes detected in the differential expression analysis were consistently regulated the same way in independent resistant lines that were not part of the comparative analysis. Hierarchical clustering of the long inverted repeat region (LIRR) gene transcription profiles grouped all susceptible samples separately from the resistant samples (S2A Fig). In particular, two blocks of genes (ostta19g00070–120 and ostta19g000560–610) were transcribed across all independent resistant lines, albeit with variation in transcription levels, but not in the susceptible controls, supporting the LIRR having a shared regulatory mechanism activated by OtV5 infection. By contrast, down-regulated genes were not consistently under-transcribed in all resistant lines (S2B Fig). This indicates over-expression of the LIRR genes was a strong determinant of resistance while under-expression of chromosome 19 genes, including the TRR, was not consistently correlated with resistance.

Altered carbohydrate synthesis and expression of diverse glycosyltransferases

Several GT genes from chromosome 19 induced in the virus-resistant state were of putative foreign origin (S2 Table) and appeared to be clustered with genes encoding functionalities related to sugar metabolism and modification (Fig 4, left gene map). These genes include a rhamnan synthesis F/Wbx GT (ostta19g00070), family 92 GT (ostta19g00600) and the co-regulated sugar modification enzymes (ostta19g00110, NAD-dependent epimerase/dehydratase; ostta19g00610, CMP-N-acetylneuraminic acid hydrolase).

Ostta19g00070 comprises merged rhamnan synthesis F and Wbx GT domains. The former is allied with bacterial rhamnose-glucose polysaccharide F (RgpF), which in *Streptococcus* transfers rhamnose to the nascent rhamnan backbone, which is incorporated into surface lipopolysaccharide [44]. RgpF domain in *O. tauri* is currently the only occurrence in eukaryotes, the species distribution being otherwise restricted to bacteria (Pfam: PF05045). Similarly, Wbx GT domain is found in bacterial gene clusters involved in synthesis of O-antigen, which in *Shigella* comprises repeated monomers of L-rhamnose, D-galacturonic acid and N-acetylgalactosamine residues [45,46]. One putative gene cluster (ostta19g00110–140) comprises a predicted NAD-dependent epimerase/dehydratase, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase, integral membrane galactosyltransferase and triose-phosphate transporter. Strikingly, the 3-beta-galactosyltransferase was associated with sequences from metazoans, while the adjacent genes appear to have homologues in the green lineage. A second putative gene cluster was a CAZy (carbohydrate-active enzymes database) family 92 GT containing a triose phosphate transporter domain (ostta19g00600) adjacent to CMP-N-acetylneuraminic acid hydrolase (ostta19g00610) of metazoan origin. These two genes were putatively involved in the activation of sugars for the synthesis of cell surface sialic acid. In proximity to these clusters of sugar modification genes was a predicted FkbM methyltransferase (ostta19g00560), which suggests it has a role in glycan methylation. Two of the up-regulated GTs (ostta19g00130 and ostta19g00630), were predicted to be membrane-associated and could be part of the Golgi mannosyltransferase complex.

Two different GTs and a triose-phosphate transporter on chromosome 19 appear to be recently generated paralogous copies because they share up to 97% amino acid identity (S3 Fig). The paralogues curiously occur in distinct genetic contexts (Fig 4). For example, the previously mentioned 3-beta-galactosyltransferase (ostta19g00120) is adjacent to sugar modification genes and is over-transcribed in resistant lines while the putative paralogue (ostta19g00320) is

adjacent to a conserved eukaryotic algae hypothetical protein and was under-transcribed (Fig 4, GT a). In this case, it suggests the same, or similar, base sugar substrate is utilised in both conditions, but alternatively modified in virus resistant lines. How these GTs and triose-phosphate transporters have been duplicated and shifted in this modular fashion is unclear, but the transposon-related genes on chromosome 19 are candidates for mediating these transfers.

The only two GTs that were differentially regulated in resistant lines that were not located on the SOC, were located on chromosome 2. These genes were down-regulated and included glycosyltransferase AER61 (ostta02g00040) and a membrane glycosyltransferase (ostta02g02060). However, other carbohydrate modification and metabolism genes located on chromosome 2 that were up-regulated in resistant lines included a somatomedin B domain-containing protein (ostta02g03940) and carbohydrate glycoside hydrolase (ostta02g04570). The former is involved in binding polysaccharides and the latter in the lysis of O-glycosidic bonds. This implicates a role also for the BOC in carbohydrate modification during viral resistance.

Transcription of OtV5

As almost all lines were R^P at the time of RNA sequencing, we explored the transcriptomes in sequenced samples for evidence of OtV5 transcription (S1 Table). In susceptible controls, a low number of read counts were assigned to few OtV5 genes (S4 Fig); in particular, OtV5_154c and OtV5_159, which correspond to highly transcribed viral specific genes (S3 Table). As susceptible cells had not been exposed to OtV5, apparent transcription of viral genes were taken to be artefacts probably arising from mis-assigned sequencing sample indexes to highly expressed transcripts within the shared sequencing flow cell lane (all samples were multiplexed together). BLAST searches at neither nucleotide levels nor protein levels revealed similarities that would intimate horizontal transfer of these genes between host and virus. As expected, lines that were R^{NP} had no OtV5 transcription and clustering with susceptible controls in their OtV5 transcription profiles (S4 Fig). The R^P lines fell into three clusters corresponding to relatively high (five samples), moderate (five samples) and negligible OtV5 transcription (seven samples near-identical to controls) showing that for most R^P lines OtV5 transcripts were nearly undetectable.

As some OtV5 genes appeared to be highly transcribed in certain R^P lines, we sought to estimate the amount of virus transcription relative to that of the host. The ratio of the mean transcript counts of the most expressed viral gene (OtV5_154c: 5,878) over that of one of the most highly expressed host genes, the ribulose biphosphate carboxylase small subunit (*rbcS*, ostta18g01880: 49,249), gives 0.12. As a comparison, the transcript abundance ratio of the corresponding genes during *Paramecium bursaria* chlorella Virus 1 (PBCV-1) infection of *Chlorella* was 37, far exceeding host transcription levels [47]. Since relative OtV5 transcription was lower than that of the host, even in R^P lines with relatively higher OtV5 transcripts levels, this suggests viral activity was certainly much lower than expected during infection of susceptible *O. tauri*.

Nonetheless, the OtV5 genes that were transcribed to a significant level (mean normalised count > 9) covered over 40% of all OtV5 genes. Expressed OtV5 genes were distributed along the entire viral genome with the expression profile varying for each sample (S5 Fig). Although we could not compare the OtV5 expression in R^P with control cells in a normal lytic cycle to confirm this, the OtV5 transcription profiles did not indicate specific genes linked to chronic infection, nor a particular stage in viral replication. Crucially, transcripts of capsid proteins were detected (S3 Table) consistent with OtV5 forming virions. This corroborates with results of the virus production assay, which detected infective lytic virions in the medium of R^P lines

(Fig 2). Apart from virion structure, transcribed OtV5 genes were associated with DNA replication, transcription, amino acid synthesis and carbohydrate modification and metabolism (S3 Table).

Expressed OtV5 carbohydrate metabolism genes included two enzymes involved in the biosynthesis of nucleotide sugars, OtV5_011, a predicted GDP-D-mannose 4,6-dehydratase, and OtV5_042, a putative dTDP-D-glucose 4,6-dehydratase. The homologue of this GDP-D-mannose 4,6-dehydratase in PBCV-1 has a dual functionality acting in both the synthesis of GDP-D-rhamnose and GDP-L-fucose, both of which are monosaccharides present in the capsid glycan but are rare in the host [48–50]. Adjacent to the GDP-D-mannose 4,6-dehydratase is a group 1 GT (OtV5_012c) that was also expressed. The spatial proximity of these genes suggests the glycosyldonor of the group 1 GT is the nucleotide sugar produced by the GDP-D-mannose 4,6-dehydratase. Three additional OtV5 GTs were transcribed including a predicted membrane-localised group 34 GT (OtV5_033), a group 2 diphosphosugar GT (OtV5_035) and a GT with no assigned CAZy group (OtV5_160). All OtV5 encoded GTs had their closest homologues in prasinoviruses indicating they have conserved virus-specific functions.

Structural rearrangements of chromosome 19 in resistant lines

Given the massive transcriptional changes in chromosome 19 and the known plasticity in outlier chromosome size between *O. tauri* strains [17], we examined the karyotypes of all experimental lines by pulsed-field gel electrophoresis (PFGE). Resistant lines showed karyotype changes, most notably as a shift in the size of chromosome 19 (34 of 36 resistant lines tested) while no change was evident in the susceptible controls (Fig 5). The most common changes were an increase of ~20–490 kb (17 lines) or a decrease of ~40–140 kb (13 lines) in the size of chromosome 19. This was likely due in the former case to duplications within chromosome 19 and in the latter, to deletions within, or fission of, chromosome 19. Several resistant lines (R1, R22, R29 and R30) appeared to have lost chromosome 19 or its location was ambiguous. In R9a and R2, there were possible translocations of regions of chromosome 19 to chromosomes 6 and 8 (respectively) or an increase in the size of chromosome 19 with concurrent insertions-deletions in other chromosomes. Indeed, variations in the size of chromosomes other than chromosome 19 were noted in R11a, R3, R5, R22, R23 and R30. More complex changes were also evident, most notably in R30 where bands apparently larger than 1.1 Mb are present that hybridize with the chromosome 19 probe. Interestingly, seven resistant lines (R9a, R13a, R1, R2, R12 R18 and R26) showed densely hybridizing material in the wells, almost always coinciding with a large increase in the size of chromosome 19 (>430 kb). This could correspond to large circular forms that cannot migrate in the PFGE [51], suggesting the extensive changes to chromosome 19 could occur via circular intermediates. However, the presence of circular DNA in the nucleus such as episomes has neither been previously detected nor specifically investigated in *O. tauri*.

Chronic virus production from lysis of a minority of susceptible cells

To investigate the chronic production of OtV5, electron microscopy was performed on two R^P lines. This showed a minority of cells (<0.5%) was visibly infected and/or in the middle of lysis (Fig 6) demonstrating viral reproduction proceeded by a typical lytic cycle. Nonetheless, a population crash of the resistant cultures was never observed, with the only loss of lines occurred following antibiotic treatment, so the frequency of lysis was low. Since all lines were re-cloned upon acquiring resistance (Fig 1), the minority of susceptible cells present in R^P lines were presumably resistant cells that had switched to a susceptible state.

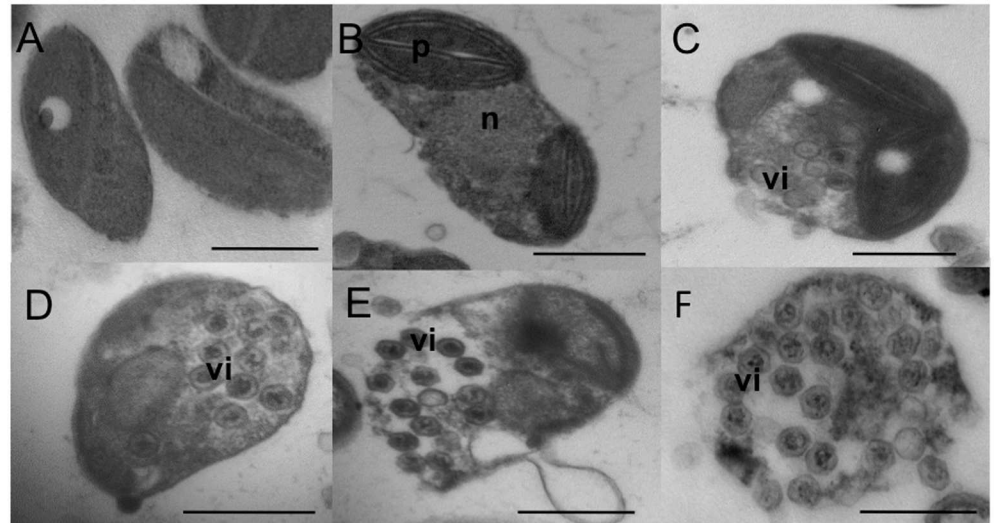


Fig 6. Virus production in two *O. tauri* resistant strains. Very few cells (less than 0.5%) show visible viral particles in their cytoplasm. Morphology of lysing R^P cells is similar to that of susceptible cells. (A) Most of the cells of *O. tauri* R^P2 strain without visible viral particles. (B) Dividing *O. tauri* R^P2 strain cell. (C) Dividing *O. tauri* R^P3 strain cell with visible intracellular viral particles (vi). (D) *O. tauri* R^P3 strain cell with visible intracellular viral particles (vi). (E) Lysis of an *O. tauri* R^P3 strain cell. (F) Lysis of an *O. tauri* R^P2 strain with visible viral particles (vi). The scale bar is 500 nm long.

doi:10.1371/journal.ppat.1005965.g006

Production of the OtV5 genome was also observed by PFGE as a ~200 kb band (Fig 5) in 15 of the 19 lines that were R^P at that time point. In the four R^P lines where the band was not apparent, viral genomes were likely present below the detection limit of DNA staining. The intensities of the viral bands were both greater and less than that of the host chromosomes suggesting, (1) a large variability in amount of viral production between R^P lines and (2), given that the number of genomes in the sample is proportional to band intensity, the number of copies of the viral genome was not directly proportional to that of the host. The fact that the OtV5 genome copy number could be much lower than that of the host implies not every cell in the host population was carrying the virus genome. Higher OtV5 band intensities could correspond to viral genomes in virions as well as multiple viral genomes generated during replication in the cell.

Similar gene functions are conserved in the SOC of all Mamiellales

Outlier chromosomes are found in all Mamiellales sequenced to date, and all SOC encode genes with a similar set of predicted functions [7], although these genes have very little detectable sequence homology compared to the non-outlier chromosomes. Only 11% of genes on the SOC of *O. lucimarinus* and *O. tauri* are orthologues, whereas the average proportion of orthologous genes is 86% on standard chromosomes [52]. We re-analysed the functional genes in all available Mamiellales genomes observing a consistent over-representation of genes involved in carbohydrate transport, synthesis and modification, particularly GTs (Table 1). Prasinoviruses are known to infect all clades of *Ostreococcus*, *Micromonas*, and *Bathycoccus* so far tested [29]. Furthermore, prasinovirus resistance has also been generated in *Bathycoccus prasinos* against BpV2 (*Bathycoccus prasinos* Virus 2) and *Micromonas pusilla* against MpV1 (*Micromonas pusilla* Virus1) in culture [28]. Given that differentially expressed genes on the SOC in *O. tauri* was strongly linked to a switch to a virus-immune state, we hypothesize the SOC has a similar role in virus resistance in these other genera of the Mamiellales.

Table 1. General characteristics of small outlier chromosomes (SOC) in Mamiellales.

	<i>O. lucimarinus</i>	<i>O. tauri</i>	<i>O. RCC809</i>	<i>B. prasinos</i>	<i>M. pusilla</i> CCMP1545	<i>M. RCC299</i>
Chromosome no.	18	19	18	19	19	17
SOC size (kb)†	150	290	102	146	242	195
ORFs	97	101	57	70	125	96
SOC gene density (gene/kb)	0.65	0.35	0.56	0.48	0.52	0.49
Global genome density (gene/kb)	0.58	0.62	0.56	0.52	0.48	0.49
SOC unknown*	49 (51%)	52 (48%)	16 (28%)	39 (56%)	71 (57%)	50 (52%)
Methyltransferase**	11	8	0	7	8	9
Glycosyltransferase	8	14	7	13	5	9
NAD epimerase/dehydratase	6	2	4	2	1	1
Sugar transporter	4	6	2	0	3	3
Other sugar metabolism***	5	10	4	4	0	3
Pyridoxal/vitB6	1	3	0	0	0	0

†Size estimate from pulsed-field electrophoresis.

*Including hypothetical (genes having similarity with other Mamiellales SOC genes) and ORFan genes.

**Including protein and/or DNA methyltransferases.

***Including lipopolysaccharide.

doi:10.1371/journal.ppat.1005965.t001

Discussion

Spontaneous mutation is not enough to explain rapid acquisition of viral resistance

Initially, we designed the experimental strategy to test whether resistant cells may arise spontaneously by mutation, and we planned to look for single nucleotide changes linked to resistance. However, the first visible signs of re-growth of resistant cells appeared about a week after inoculation in all of the cultures. Since cultures do not become visibly green until they reach at least 10^6 cells.ml⁻¹, and assuming the cells have an optimum division rate under these conditions (1.4 divisions.day⁻¹ in this growth chamber, M. Krasovec, personal communication), we find that about 1 in 1,000 cells may have become resistant at the time of or just after inoculation with the virus. This, however, may be an underestimation, since after a shock (such as after sub-culturing), there is usually a lag before re-growth. This frequency far exceeds the expected spontaneous mutation rate in *O. tauri* since comparison of the re-sequenced *O. tauri* genome between 2001 and 2009 revealed the fixation of eight single nucleotide substitutions and two deletions during the approximately 6,000 generations in the lab [4]. We thus preferred an alternative hypothesis, that resistance is induced by the biotic challenge of virus infection, for example by epigenetic modifications affecting gene expression patterns. Since the original culture was clonal, it seemed unlikely that a proportion of host cells was resistant at the moment of inoculation; although we cannot rule out that a reversible regulatory switch to a resistant state occurred in a small proportion of cells analogous to the minority of cells in R^P lines that was lysed by OtV5. However, after a period of selection, resistance in the majority of cells appears to be stable since none of the resistant lines were lysed after re-infection over the course of the experiment.

Extensive alteration of glycosylation state as the principal driver of virus immunity

In several loci, up-regulated genes on the LIRR related to carbohydrate metabolism and modification appeared to be spatially grouped. We speculate genes that are clustered together act on

the same carbohydrate synthesis and glycosylation pathway. Notably, genes uniquely expressed in virus-resistant lines might dramatically alter the saccharide composition and glycosylation state of the cell. Several of the highly expressed GTs are associated with synthesis of surface glycans. Surface glycans are known to be important for host–virus interactions, where they mediate initial binding and recognition events of both host cells and pathogens at the cell surfaces [53]. In particular, rhamnan is a polymer important for bacterial host–virus interactions [54,55], and sulphated derivatives of rhamnans are known to have anti-viral activity in mammals [56]. The metazoan-derived CMP N-acetylneuraminic acid hydroxylase (ostta19g00610), in humans makes the influenza A virus receptor, N-acetylneuraminic acid, whose decorations affect receptor specificity [57]. Likewise, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase (ostta19g00120), known to function in the synthesis of extended mucin type O-linked glycans, has been associated with certain immune mediated diseases in humans [58]. Thus, expression of these genes may affect resistance through masking, or altering the *O. tauri* surface receptor used by OtV5 to inhibit viral adsorption. However, Thomas et al., [28] observed no statistical difference in OtV5 adsorption to R^P, R^{NP} and susceptible cells suggesting defective adsorption was not the mechanism for viral resistance. This may be due in part the very low proportion of *O. tauri* cells ‘competent’ to OtV5 adsorption at a given time point; even after inoculating susceptible *O. tauri* with high titres of OtV5, at most 20% of cells had visibly adsorbed virions despite the fact the majority of cells subsequently lyse [24]. One possible explanation is that the OtV5 receptor is only available during a defined point in host cell cycle, and as *O. tauri* cells are not perfectly synchronised in culture [59], adsorption is staggered over time making differences difficult to detect by a standard adsorption assay.

Alternatively, the over-expression of these specific host GTs may curtail OtV5 replication at another stage of the replication cycle, one possible candidate being virion assembly. In the model Chlorovirus, PBCV-1, the major capsid protein is glycosylated at least six sites and the glycans include rare monosaccharides whose biosynthesis genes are virus-encoded (see Piacente et al., [60] and references therein). As OtV5 similarly encodes numerous GTs and homologues to the PBCV-1 sugar metabolism enzymes, many of which are conserved in prasinoviruses (S3 Table), glycoconjugates are likely to be similarly essential to virion structure. Viral resistance may arise by the host switching carbohydrate metabolic pathways, thereby perturbing proper virion formation. This could be mediated, for instance, by feedback inhibition of viral sugar biosynthesis pathways by host generated products, directly altering the viral glycosyldonors/acceptors or simply altering the substrate pool. For example, virally produced GDP-D-rhamnose may be depleted by the host by redirecting it into extracellular rhamnan. Much remains unknown about the glycobiology of both prasinoviruses and Mamiellales and studies of their sugar composition, glycan structure and functional characterisation of GTs will help to pinpoint how the host–virus interaction is suppressed during resistance.

The link between resistance to prasinoviruses and alterations of glycosylation might extend to other Mamiellales, since there are numerous sugar modification genes clustered on the SOC in other members of the Mamiellales with sequenced genomes. They are extremely diverse, with different GT genes in each species, in keeping with the notion that they are important in viral specificity.

Occasional viral lysis *via* low-level instability in the resistance phenotype?

Thomas et al. [28] reported that chronically infected cells appeared to release viruses by budding, but after examination of many electron micrographs, we found that viral-induced lysis of host cells was occurring, which was further supported by the overall low level of OtV5

transcription in R^P lines. R^P cell lines are more likely to be unstably resistant, giving rise to a small proportion of revertant susceptible cells. Thus, apparent chronic virus production occurs by lysis of a minority of susceptible cells emerging in a majority of uninfected population at a frequency sufficient to propagate OtV5 over subsequent generations, although the serial cell transfers during the experiment tended towards eventual OtV5 extinction. This further suggests the two apparent kinds of resistance, R^P and R^{NP} , may reflect observing virus production at the level of a batch culture population and does not correspond to distinct mechanisms of immunity. Over time, revertant susceptible cells arising in cultures where OtV5 were absent could competitively exclude resistant strains, although this was not observed within the almost two-year duration of the study. In keeping with this observation R^P cells have been shown to have growth rates similar to wild-type in culture, and differences in fitness could only be shown by competition between co-cultured strains [28].

Genetic plasticity and transcriptional regulation of the SOC

It is tempting to speculate that the biotic stress induced, by an unknown mechanism, the activation of a transposon that mediated rearrangements in chromosome 19. Activation of LTR-retrotransposons during pathogen attack was observed in tobacco [61] and has since been observed following diverse kinds of biotic or abiotic stresses in diverse organisms (reviewed in [62–64]). This process may lie at the heart of genomic plasticity in response to environmental changes, since genome evolution might be facilitated by bursts of transposon activity [65]. In eukaryotic marine plankton, retrotransposons are probably the most abundantly expressed sequences [66]. The LTR of retrotransposons carry promoters that can affect the transcription of adjacent genes, although the effects of environmental conditions on LTR retrotransposons, which require transcriptional activity for their activation, are complex [62,67,68]. The putative TRIM on chromosome 19 (S1 Fig) may be activated in *trans* by the predicted complete version of this element, *retrostreo2*, on chromosome 8. However, since its reverse transcriptase (*osta08g00390*) was down-regulated in resistant lines (S4 Table), it is not a likely candidate for causing changes in the SOC sequence. In addition, the tandem repeats surrounding the retroelement do not look like LTRs, they may be too long (2.3 kb, S1 Fig) and it seems unlikely that transcription of the TRR could influence the LIRR and the other distant and extensive set of resistance genes.

In Mamiellales, DNMT5 was proposed to function as the CpG maintenance methyltransferase [69] as DNMT1 and DNMT3 are lacking, the latter two enzymes being responsible for both *de novo* and maintenance CpG methylation in other organisms [70]. However, extensive CpG methylation occurs in nucleosome linkers and thus apparently functions to guide nucleosome positioning within the extremely compact nucleus. This raises the question of whether or not CpG methylation still plays a role in gene silencing in picoeukaryotes, albeit mediated by an alternative DNMT. A possible MT involved in epigenetic silencing is *osta19g00045*, a predicted 5-cytosine methyltransferase located on chromosome 19 that was over-transcribed in the resistant lines (S2 Table). However, if indeed silencing of the LIRR of chromosome 19 in the susceptible control cells occurs by CpG methylation as in plants and animals [70], we would expect decreased activity of *de novo* CpG MT and increased activity of demethylases, which was not the case in this study. This suggests that silencing of chromosome 19 occurs *via* an alternative mechanism.

A more likely candidate is an up-regulated DDE superfamily endonuclease (Fig 4) shared between both outlier chromosomes (*osta19g00165* and *osta02g00275*) that may mediate rearrangements. Type II transposons can carry silencer domains in the C-terminal region of their transposases [71] that not only repress their own expression, but can repress expression from

adjacent genes, probably by histone modifications that change chromatin conformation. Expression of these transposons can be induced by certain transcription factors. Given that this gene lies at the heart of the over-transcribed region, it might explain both the transcriptional switch and the genome instabilities that we observe. However, it is also possible that an unknown mechanism up-regulating the inverted repeat region leads to de-repression of the predicted endonuclease, leading to the genomic instability observed. In *O. lucimarinus* both of the predicted orthologues of *osta19g00165* reside on the chromosome 2, but we have not yet examined the karyotypes of this species after viral infection. Three predicted transcription factors are up-regulated in resistant lines, *osta15g02170*, containing a Myb-SANT domain of the REB1 family [72]; *osta01g00380*, whose N-terminal regions shows weak similarity to RNA polymerase III subunit delta (BLAST to *pfam12657*, 1.39e-06) that is important in synthesis of some essential non-coding RNAs [73]; and *osta08g00865*. Further experimental investigation would be required to determine whether they play a role in transcriptional activation of the SOC, since these conserved domains are important in numerous kinds of regulatory processes. We might thus speculate that one of these up-regulated transcription factors bind to the *osta19g00165* C-terminal domain after viral attack, inducing the adjacent glycosyltransferases, perhaps by changing chromatin configuration, and activating this transposon. Since the region is a large inverted repeat, an endonuclease target site might also be duplicated and physically distant, perhaps explaining the large size variations observed in this chromosome between independent resistant lines.

Given the presence of numerous repeated regions, and the reported occurrence of homologous recombination in *O. tauri* [12], a further alternative might be that the selective pressure for resistance to viruses promotes the growth of relatively rare cells in which genomic rearrangements have arisen by recombination, leading to changed genomic contexts and activation of the inverted repeat region. We have, however, never observed a change in the size of the *O. tauri* SOC in culture in our routine laboratory strain RCC4221 over a period of about 20 years at ~300 kb ([5,17] and this work). We were surprised to discover that extensive karyotypic changes had occurred in practically all of resistant lines entailing both increase and decrease in size. This suggests insertions/deletions of genetic material to the SOC after the acquisition of resistance to OtV5 is a consequence of activating resistance genes and not necessarily causative. Future work examining the SOC sequence in virus resistant *O. tauri* strains as well as in strains isolated from the environment and monitoring SOC changes during the critical period when resistance is established would inform which genetic changes occur and their relationship to viral resistance.

Parallels to other host–pathogen systems and evolution of a viral immunity chromosome

Several parallels with other genetic systems that involve biotic interactions between partners of symbionts or pathogens have come to light in our system. In the marine cyanobacterium *Prochlorococcus*, genes involved in resistance to viruses are clustered in hypervariable genomic islands [74] that facilitate rapid evolution of resistance and persistence of the host in some species [75]. As in *Ostreococcus*, in these regions most of these genes may be involved in cell surface interactions. Supernumerary chromosomes, which can vary in size and number between members of the same biological species, have been found in diverse fungal pathogens (reviewed in [76]). They often differ in base composition and structure with other autosomes of the same species, and they can carry genes encoding pathogenicity effectors, probably facilitating the rapid evolution of virulence. In *Nectria haematococca*, for example, pathogenicity effectors are clustered on a dispensable chromosome with an altered GC content that also carries

transposons [77] and in *Fusarium* such chromosomes can be transferred horizontally between strains, broadening their host specificity [78]. In *Ostreococcus*, the tables are turned, since it is the host cell that shows genomic plasticity, whereas prasinovirus genomes show less nucleotide variation between strains than those of their host genomes [26].

We speculate that the evolution of the small outlier chromosome as a viral immunity chromosome in the Mamiellales may have arisen following selection for small genome size, since genome compaction probably led to loss of genes involved in the silencing of viral attack by small RNAs, subsequently forcing the evolution of the defence mechanism typical to this group of algae. Small RNAs are central for epigenetic modifications and genome stability in eukaryotes in general [79,80], and in higher plants mechanisms of resistance to viral infection have been widely studied, where small RNAs have a central role in controlling viral infections (see [77] for a review). Although *Ostreococcus* is in the green lineage, most of the analysed genomes of algae in this group lack detectable homologues of the canonical genes *Argonaut* and *Dicer* thought to be important for their fabrication [81], possibly because it lost these genes during evolution of its extremely small genome that may result from “streamlining” in its marine environment [82]. Loss or gain of the biochemical machinery for fabrication of different kinds of small RNA has led to diversification in eukaryotic systems, and in some well-studied fungal systems loss of gene functions might provide clues about alternative pathways for dealing with viral assault [83].

Conclusions

To our knowledge, this is the first time that a chromosome specialized in defence against viruses has been described. Our results throw some light on the variability in the size of the SOC observed between different wild-type isolates of the *O. tauri* population in coastal north-western Mediterranean Sea [17]. These waters are indeed infested with prasinoviruses [84], and prasinoviruses are similarly abundant in the English Channel, where they infect *Micromonas* spp. [85]. While the precise mechanism for resistance remains to be elucidated by subsequent experimental investigations, we show that challenging cultures with a compatible virus reliably leads to extensive chromosome rearrangements, providing an opportunity to study natural genome evolution in controlled laboratory conditions in *O. tauri*. The observed rearrangements in the *O. tauri* SOC may be produced by an adaptive mechanism that quickly generates genetic variants able to evade viral lysis. Clustering of the genetic features observed in the SOC helps to explain several of its unusual characteristics. In particular, a faster rate of evolution than the other chromosomes, evidenced in the high level of species-specific genes. The conservation of the same functional gene categories, but not of gene homologues, would also be consistent with rapid “arms-race” evolution against the particular prasinoviruses that infect each species and implies the SOC as a whole has a conserved function in viral immunity.

Materials and Methods

Culture conditions and experimental design

The strains used in the experiment were *O. tauri* RCC4221 (Roscoff Culture Collection [86] and *Ostreococcus tauri* virus 5 (OtV5). Liquid L1 medium (NCMA, Bigelow Laboratory for Ocean Sciences, USA) was used for culturing throughout and prepared using autoclaved seawater from offshore Banyuls Bay (MOLA station: 42°27'11"N, 3°8'42"E) diluted 10% with MilliQ water and filtered prior to use through 0.22 μm filters. All cultures were maintained under a 12:12 hour light/dark regime in 100 $\mu\text{mol photon m}^{-2} \text{s}^{-1}$ white light at 20°C. To grow *O. tauri* on a solid medium, molten agarose (1.5%) equilibrated to 60°C was added to cultures to give a final agarose concentration: 0.15% and immediately poured into 9 cm diameter petri

dishes. Single *O. tauri* colonies were picked from the solid medium using a sterile pipette tip and placed into liquid L1. OtV5 plaques were produced by mixing serial dilutions of OtV5 lysate with *O. tauri* RCC4221 culture and growing on solid medium [24]. A stock OtV5 lysate used in all subsequent infections was produced by randomly selecting a plaque, inoculating it into an *O. tauri* RCC4221 culture, purifying the resulting lysate (centrifugation at 8,000 g for 20 min and filtration through 0.2 μm) and storing at 4°C.

Viral susceptibility and virus production tests

To test viral susceptibility, stock OtV5 lysate was added to *O. tauri* lines growing in 48-well plates with the precursor susceptible *O. tauri* RCC4221 serving as a positive control. Lines were considered susceptible if visible lysis occurred within approximately one week. To test if *O. tauri* resistant lines were producing infective virus, an aliquot of the culture medium from actively growing cells was sampled, then centrifuged at 8,000 g for 20 min to remove *O. tauri* cells. The cell-free medium was added to susceptible *O. tauri* RCC4221 growing in 48-well plates. *O. tauri* lines were considered to be producing virus if visible lysis occurred in the susceptible line within approximately one week.

Nucleic acid extraction, sequencing and assembly

DNA was extracted from the initial 1 L culture of the *O. tauri* from which all subsequent lines were cloned (Fig 1) using a CTAB extraction method [34]. GATC Biotech performed whole genome sequencing using six single molecule real time (SMRT) cells on the PacBio RS II platform and *de novo* assembly with the Hierarchical Genome Assembly Process (HGAP, Pacific Biosciences). The final updated chromosome 19 sequence was obtained by scaffolding the HGAP-generated unitigs in Geneious [87,88]. Before total RNA extraction, *O. tauri* culture lines were subjected to antibiotic treatment to reduce the amount of bacteria in the cultures according to [89]. The antibiotic treated culture was used to inoculate a 200 mL culture for RNA extraction. Cultures were grown to exponential phase (11–32 million cells mL^{-1}) and cells harvested by centrifugation at 8,000 g for 20 min in a swinging bucket rotor during the first 3 hours of the light cycle. Total RNA was extracted using the Direct-zol RNA kit (Zymo Research). Cell densities for each culture line are recorded in S1 Table. Selection for polyadenylated RNA, library preparation and sequencing was performed at GeT (Génome et Transcription, GENOTOUL, Toulouse, France). RNA libraries were sequenced on the Illumina Hi-Seq 2000 platform by multiplexing all samples on a single flowcell lane, which generating paired end reads of 101 bp in length. RNA sequence reads were checked for quality using FastQC [90].

Differential gene transcription analysis

In order to have a balanced experimental design and to control for batch effects, four OtV5-resistant and four control lines that were from matching RNA extraction batches were selected for differential gene expression analysis (S1 Table). Transcriptome read pairs (fragments) were aligned using TopHat2 [91] (alignment parameters: -i 10 -I 4000, -G) to the annotated genome sequence of *O. tauri* RCC4221 [4]. The counts of fragments aligning to each gene was determined using the htseq-count function of HTSeq [92] with parameters (-t CDS -s no -m intersection-nonempty). Exploratory and differential gene expression analyses were performed on fragment count tables using the R package DESeq2 [93]. Hierarchical clustering of the distance between transcriptome profiles after correction for batch effects separated resistant and susceptible lines in two groups (S5 Fig). For differential gene expression, RNA batch number was used as a secondary design variable to adjust for batch effects and the primary variable tested

was virus resistance vs susceptibility using the DESeq function accepting genes as significantly differentially transcribed with adjusted p-value <0.1.

Pulsed-field gel electrophoresis and in-gel hybridization

PFGE and in-gel hybridization was conducted as previously described [94]. Briefly, cell cultures were grown to mid-exponential phase, harvested by centrifugation at 8000 g for 20 min and cells were resuspended in TE buffer (10 mM Tris-HCl, 125mM EDTA, pH 8) at a density of 8.7×10^8 cells ml⁻¹. Electrophoresis was performed in 0.8% agarose gels in 0.5× TBE buffer (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA at pH 8) using the CHEF-DR III (Bio-Rad) system. For each sample, 2 mm of plug was loaded into the wells. Electrophoresis was run at 6 V cm⁻¹ at 14°C with 120° pulse angle for 15 h with a switch time of 60 s and followed 9 h at a switch time of 90 s. After PFGE, DNA was chemically denatured and dehydrated in a vacuum dryer. PCR was used to amplify a 1681 bp region of the *O. tauri* gene osta19g00640 (5052 bp), located on chromosome 19, with the primer pair 19e_Fw 5'-GCGATGCGGTGCTCTACC-3' and 19e_Rv 5'-CGTGGAGTTATCCCCGAACC-3' (PCR conditions described in [17]). This gene was chosen as a probe for chromosome 19 because it was consistently transcribed between resistant and susceptible lines (S6 Fig), and thus was not deleted and it has a putative orthologue in the SOC of *O. lucimarinus* suggesting a conserved function. After gel purification of the PCR product (Wizard-Prep, kit, Machery-Nagel), the amplicon was randomly labelled with [α -³²P]CTP (Perkin-Elmer) according to manufacturer's instructions (Prime-a-Gene kit, Promega) for use as a DNA probe. Dried gels were equilibrated in hybridization buffer (6× SSC, 5× Denhardt's solution, 0.1% (w/v) sodium dodecyl sulphate, 10 μ g ml⁻¹ tRNA), radiolabelled probe was added, hybridized overnight at 65°C and the gel exposed to radiographic film.

Accession numbers

Sequence data used in this study can be found in the GenBank data libraries. *O. tauri* RCC4221 chromosome sequences are under accession numbers CAID01000001.2 to CAID01000020.2 [4] except for the updated chromosome 19 sequence (this study), which is available from <http://www.php.obs-banyuls.fr/publications/data/2/>. Transcriptome data is available from PRJNA344946. The updated genome sequence and annotation of OtV5 [24] is under accession EU304328.2. Gene models and annotations for *O. tauri* RCC4221 are also available from the Online Resource for Community Annotation of Eukaryotes (ORCAE) under *Ostreococcus tauri* V2 [95, 96].

Transmission electron microscopy

Transmission electron microscopy was performed as previously described in [24] with some modifications. Briefly, 200 mL cultures of *O. tauri* in exponential phase were fixed in 1% glutaraldehyde for 30 min and fixed cells were centrifuged for 30 min at 2500 g and the cell pellet quickly resuspended in molten (37°C) 1% low melting point agarose (agarose type II, Sigma) and allowed to set in a disposable micropipette (SMI, Emerville, CA, USA). The agarose cell plug was fixed in buffer of one volume of sodium cacodylate (0.4M) in two volumes of L1 medium containing 2.5% glutaraldehyde for 2 h, then washed in the same buffer without glutaraldehyde. Post fixation was performed in 1% OsO₄ in sodium cacodylate (0.2M) for 1 h. After two 15 min washes in sodium cacodylate (0.2M), the agarose cell plug was cut into small pieces and dehydrated in ethanol and embedded in Epon 812 resin at 60°C for 48 h. Ultra-thin slices (80–90 nm) were placed on a 300 mesh grid and stained with uranyl acetate for 15 min, followed by lead citrate staining for 2 min then visualised with a Hitachi H 7500 transmission electron microscope.

Supporting Information

S1 Fig. Organisation of the ~37 kb tandem repeat region on chromosome 19. (A) Light blue blocks represent 2250–2259 bp repeated monomers with the first and last two nucleotides of the repeat shown if present and the internal magenta arrow indicating the position of the predicted gene, *ostta19g00035*. Two repeats were interrupted, one by 3354 bp (white rectangle) containing the predicted ORF, *ostta19g00240* (black arrow) and the other by a 2232 bp putative TRIM, black bars indicate target site duplication, blue arrows indicate LTRs. (B) Zoom in of map comparing the chromosome 19 TRIM to *retrostreo2*, an LTR-retrotransposon on chromosome 8 (5910 bp). Blue arrows are LTRs, purple arrow is the putative POL (*ostta08g00390*), orange arrow is a gene of unknown function containing a predicted zinc finger domain (*ostta08g00400*), light green bar is the putative polypurine tract and yellow bar is the putative primer binding site. Red blocks joining the two genomic maps indicate blocks of nucleotide identity (99% in the LTRs, 91% in the central region). The putative genes in the TRIM are likely disrupted.

(TIFF)

S2 Fig. Heat map of differentially regulated genes on chromosome 19 in *O. tauri* samples. RNA fragment counts from alignment to the *O. tauri*, organelle and OtV5 genomes were transformed (regularised \log_2) and counts of chromosome 19 genes were used to generate the heat map. (A) Up-regulated and (B) down-regulated genes from differential gene analysis on chromosome 19 are presented in the order they occur on the genome from left to right (see Fig 4). Virus-susceptible lines are S2a–S5a (blue font), all other samples were OtV5-resistant.

(TIFF)

S3 Fig. Multiple sequence alignments of amino acid sequences of putative paralogous genes located on chromosome 19. (A) Putative beta1,3-galactosyltransferase, (B) putative galactosyltransferase, (C) putative glycosyltransferase family 25 and (D) putative triose phosphate transporter. Gene identifiers are marked on the left in blue. The positions where the alignment starts and ends are marked on either side of the sequence and the full amino acid sequence length is in parentheses.

(TIFF)

S4 Fig. Heat map of OtV5 gene transcription profiles in *O. tauri* samples. RNA fragment counts from alignment to the *O. tauri*, organelle and OtV5 genomes were transformed (regularised \log_2) and counts of OtV5 genes were used to generate the heat map. Virus-susceptible lines are S2a–S5a (blue font), all other samples were OtV5-resistant. R24 and R26 (yellow) were R^{NP} while all other resistant lines were R^P at the time of RNA sequencing.

(TIFF)

S5 Fig. Heat map of the sample-to-sample Euclidean distances of *O. tauri* transcriptomes. RNA fragment counts from alignment to the *O. tauri*, organelle and OtV5 genomes were transformed (regularised \log_2) and sample–sample distances calculated (A) without fitting and (B) after fitting for the effect of RNA processing batch. Virus susceptible lines are S2a–S5a and resistant lines are R5, R6, R16 and R13a. Sample identifiers from the same RNA processing batch have the same colour.

(TIFF)

S6 Fig. Alignment of RNA-Seq reads to the region of chromosome 19 used as a hybridisation probe. Magenta arrows indicate the positions of the forward and reverse primer pair used to generate the probe.

(TIFF)

S1 Table. *O. tauri* transcriptome sequencing data.
(DOCX)

S2 Table. Differentially transcribed genes in OtV5-resistant *O. tauri* on chromosome 19.
(DOCX)

S3 Table. OtV5 genes transcribed in resistant *O. tauri* lines.
(DOCX)

S4 Table. Differentially transcribed genes in OtV5-resistant *O. tauri* from chromosomes other than chromosome 19.
(DOCX)

Acknowledgments

We thank Marie-Line Escande, Leila Kerkache and the microscopy platform of the Oceanological Observatory of Banyuls for their contribution of the electron microscopy analysis, the GeT (Génome et Transcriptome) core facility for the transcriptomic and genomic sequencing, Benoît Piégu for assistance with annotation of transposable elements, Rozenn Thomas for preliminary observations made during her Ph.D. studies (2011, Sorbonne Universités, Paris, in French), and the Genophy (Evolutionary and environmental genomics of Phytoplankton) team for useful discussion.

Author Contributions

Conceptualization: NG HM GP.

Funding acquisition: NG.

Investigation: SY CH GP NG HM ED.

Methodology: NG HM GP.

Supervision: NG.

Writing – original draft: SY HM NG.

Writing – review & editing: GP ED.

References

1. Treusch AH, Demir-Hilton E, Vergin KL, Worden AZ, Carlson CA, Donatz MG, et al. Phytoplankton distribution patterns in the northwestern Sargasso Sea revealed by small subunit rRNA genes from plastids. *ISME J.* 2012; 6: 481–492. doi: [10.1038/ismej.2011.117](https://doi.org/10.1038/ismej.2011.117) PMID: [21955994](https://pubmed.ncbi.nlm.nih.gov/21955994/)
2. Demir-Hilton E, Sudek S, Cuvelier ML, Gentemann CL, Zehr JP, Worden AZ. Global distribution patterns of distinct clades of the photosynthetic picoeukaryote *Ostreococcus*. *ISME J.* 2011; 5: 1095–1107. doi: [10.1038/ismej.2010.209](https://doi.org/10.1038/ismej.2010.209) PMID: [21289652](https://pubmed.ncbi.nlm.nih.gov/21289652/)
3. Vault D, Eikrem W, Viprey M, Moreau H. The diversity of small eukaryotic phytoplankton (< 3 μm) in marine ecosystems. *Fems Microbiology Reviews.* 2008; 32: 795–820. doi: [10.1111/j.1574-6976.2008.00121.x](https://doi.org/10.1111/j.1574-6976.2008.00121.x) PMID: [18564290](https://pubmed.ncbi.nlm.nih.gov/18564290/)
4. Blanc-Mathieu R, Verhelst B, Derelle E, Rombauts S, Bouget F-Y, Carré I, et al. An improved genome of the model marine alga *Ostreococcus tauri* unfolds by assessing Illumina de novo assemblies. *BMC Genomics.* 2014; 15: 1103. doi: [10.1186/1471-2164-15-1103](https://doi.org/10.1186/1471-2164-15-1103) PMID: [25494611](https://pubmed.ncbi.nlm.nih.gov/25494611/)
5. Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, Robbens S, et al. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proceedings of the National Academy of Sciences of the United States of America.* 2006; 103: 11647–11652. doi: [10.1073/pnas.0604795103](https://doi.org/10.1073/pnas.0604795103) PMID: [16868079](https://pubmed.ncbi.nlm.nih.gov/16868079/)

6. Palenik B, Grimwood J, Aerts A, Rouze P, Salamov A, Putnam N, et al. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104: 7705–7710. doi: [10.1073/pnas.0611046104](https://doi.org/10.1073/pnas.0611046104) PMID: [17460045](https://pubmed.ncbi.nlm.nih.gov/17460045/)
7. Moreau H, Verhelst B, Couloux A, Derelle E, Rombauts S, Grimsley N, et al. Gene functionalities and genome structure in *Bathycoccus prasinos* reflect cellular specializations at the base of the green lineage. *Genome Biology*. 2012; 13: R74. doi: [10.1186/gb-2012-13-8-r74](https://doi.org/10.1186/gb-2012-13-8-r74) PMID: [22925495](https://pubmed.ncbi.nlm.nih.gov/22925495/)
8. Worden A, Lee J-H, Mock T, Rouzé P, Simmons MP, Aerts AL, et al. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science*. 2009; 324: 268–272. doi: [10.1126/science.1167222](https://doi.org/10.1126/science.1167222) PMID: [19359590](https://pubmed.ncbi.nlm.nih.gov/19359590/)
9. Aylward FO, Eppley JM, Smith JM, Chavez FP, Scholin CA, DeLong EF. Microbial community transcriptional networks are conserved in three domains at ocean basin scales. *PNAS*. 2015; 112: 5443–5448. doi: [10.1073/pnas.1502883112](https://doi.org/10.1073/pnas.1502883112) PMID: [25775583](https://pubmed.ncbi.nlm.nih.gov/25775583/)
10. Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA, et al. Pattern and synchrony of gene expression among sympatric marine microbial populations. *PNAS*. 2013; 110: E488–E497. doi: [10.1073/pnas.1222099110](https://doi.org/10.1073/pnas.1222099110) PMID: [23345438](https://pubmed.ncbi.nlm.nih.gov/23345438/)
11. Vaultot D, Lepère C, Toulza E, De la Iglesia R, Poulain J, Gaboyer F, et al. Metagenomes of the Picoalga *Bathycoccus* from the Chile Coastal Upwelling. *PLoS ONE*. 2012; 7: e39648. doi: [10.1371/journal.pone.0039648](https://doi.org/10.1371/journal.pone.0039648) PMID: [22745802](https://pubmed.ncbi.nlm.nih.gov/22745802/)
12. Lozano J-C, Schatt P, Botebol H, Vergé V, Lesuisse E, Blain S, et al. Efficient gene targeting and removal of foreign DNA by homologous recombination in the picoeukaryote *Ostreococcus*. *Plant J*. 2014; 78: 1073–1083. doi: [10.1111/tpj.12530](https://doi.org/10.1111/tpj.12530) PMID: [24698018](https://pubmed.ncbi.nlm.nih.gov/24698018/)
13. Vandepoele K, Van Bel M, Richard G, Van Landeghem S, Verhelst B, Moreau H, et al. pico-PLAZA, a genome database of microbial photosynthetic eukaryotes. *Environmental Microbiology*. 2013; 15: 2147–2153. doi: [10.1111/1462-2920.12174](https://doi.org/10.1111/1462-2920.12174) PMID: [23826978](https://pubmed.ncbi.nlm.nih.gov/23826978/)
14. Piganeau G, Grimsley N, Moreau H. Genome diversity in the smallest marine photosynthetic eukaryotes. *Research in Microbiology*. 2011; 162: 570–577. doi: [10.1016/j.resmic.2011.04.005](https://doi.org/10.1016/j.resmic.2011.04.005) PMID: [21540104](https://pubmed.ncbi.nlm.nih.gov/21540104/)
15. Li X-Q, Du D. Variation, Evolution, and Correlation Analysis of C+G Content and Genome or Chromosome Size in Different Kingdoms and Phyla. *PLoS ONE*. 2014; 9: e88339. doi: [10.1371/journal.pone.0088339](https://doi.org/10.1371/journal.pone.0088339) PMID: [24551092](https://pubmed.ncbi.nlm.nih.gov/24551092/)
16. Grimsley N, Yau S, Piganeau G, Moreau H. Typical Features of Genomes in the Mamiellophyceae. In: Ohtsuka S, Suzuki T, Horiguchi T, Suzuki N, Not F, editors. *Marine Protists*. Springer Japan; 2015. pp. 107–127. http://link.springer.com/chapter/10.1007/978-4-431-55130-0_6
17. Subirana L, Péquin B, Michely S, Escande M-L, Meilland J, Derelle E, et al. Morphology, Genome Plasticity, and Phylogeny in the Genus *Ostreococcus* Reveal a Cryptic Species, *O. mediterraneus* sp. nov. (Mamiellales, Mamiellophyceae). *Protist*. 2013; 164: 643–659. doi: [10.1016/j.protis.2013.06.002](https://doi.org/10.1016/j.protis.2013.06.002) PMID: [23892412](https://pubmed.ncbi.nlm.nih.gov/23892412/)
18. Bratbak G, Thingstad F, Heldal M. Viruses and the microbial loop. *Microb Ecol*. 1994; 28: 209–221. doi: [10.1007/BF00166811](https://doi.org/10.1007/BF00166811) PMID: [24186448](https://pubmed.ncbi.nlm.nih.gov/24186448/)
19. Suttle CA. Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol*. 2007; 5: 801–12. doi: [10.1038/nrmicro1750](https://doi.org/10.1038/nrmicro1750) PMID: [17853907](https://pubmed.ncbi.nlm.nih.gov/17853907/)
20. Mojica KDA, Brussaard CPD. Factors affecting virus dynamics and microbial host-virus interactions in marine environments. *FEMS Microbiol Ecol*. 2014; 89: 495–515. doi: [10.1111/1574-6941.12343](https://doi.org/10.1111/1574-6941.12343) PMID: [24754794](https://pubmed.ncbi.nlm.nih.gov/24754794/)
21. Dodds JA. Viruses of marine algae. *Experientia*. 1979; 35: 440–442. doi: [10.1007/BF01922694](https://doi.org/10.1007/BF01922694)
22. Van Etten JL, Lane LC, Dunigan DD. DNA viruses: the really big ones (giruses). *Annu Rev Microbiol*. 2010; 64: 83–99. doi: [10.1146/annurev.micro.112408.134338](https://doi.org/10.1146/annurev.micro.112408.134338) PMID: [20690825](https://pubmed.ncbi.nlm.nih.gov/20690825/)
23. Mayer JA, Taylor FJR. Virus Which Lyses the Marine Nanoflagellate *Micromonas pusilla*. *Nature*. 1979; 281: 299–301.
24. Derelle E, Ferraz C, Escande M-L, Eychenie S, Cooke R, Piganeau G, et al. Life-cycle and genome of OtV5, a large DNA virus of the pelagic marine unicellular green alga *Ostreococcus tauri*. *PLoS ONE*. 2008; 3: e2250, 1–13. doi: [10.1371/journal.pone.0002250](https://doi.org/10.1371/journal.pone.0002250) PMID: [18509524](https://pubmed.ncbi.nlm.nih.gov/18509524/)
25. Derelle E, Monier A, Cooke R, Worden AZ, Grimsley NH, Moreau H. Diversity of Viruses Infecting the Green Microalga *Ostreococcus lucimarinus*. *J Virol*. 2015; 89: 5812–5821. doi: [10.1128/JVI.00246-15](https://doi.org/10.1128/JVI.00246-15) PMID: [25787287](https://pubmed.ncbi.nlm.nih.gov/25787287/)
26. Moreau H, Piganeau G, Desdevises Y, Cooke R, Derelle E, Grimsley N. Marine Prasinovirus Genomes Show Low Evolutionary Divergence and Acquisition of Protein Metabolism Genes by Horizontal Gene Transfer. *J Virol*. 2010; 84: 12555–12563. doi: [10.1128/JVI.01123-10](https://doi.org/10.1128/JVI.01123-10) PMID: [20861243](https://pubmed.ncbi.nlm.nih.gov/20861243/)

27. Weynberg KD, Allen MJ, Gilg IC, Scanlan DJ, Wilson WH. Genome Sequence of *Ostreococcus tauri* Virus OtV-2 Throws Light on the Role of Picoeukaryote Niche Separation in the Ocean. *J Virol*. 2011; 85: 4520–4529. doi: [10.1128/JVI.02131-10](https://doi.org/10.1128/JVI.02131-10) PMID: [21289127](https://pubmed.ncbi.nlm.nih.gov/21289127/)
28. Thomas R, Grimsley N, Escande M, Subirana L, Derelle E, Moreau H. Acquisition and maintenance of resistance to viruses in eukaryotic phytoplankton populations. *Environmental Microbiology*. 2011; 13: 1412–1420. doi: [10.1111/j.1462-2920.2011.02441.x](https://doi.org/10.1111/j.1462-2920.2011.02441.x) PMID: [21392198](https://pubmed.ncbi.nlm.nih.gov/21392198/)
29. Bellec L, Clerissi C, Edern R, Foulon E, Simon N, Grimsley N, et al. Cophylogenetic interactions between marine viruses and eukaryotic picophytoplankton. *BMC Evolutionary Biology*. 2014; 14: 59. doi: [10.1186/1471-2148-14-59](https://doi.org/10.1186/1471-2148-14-59) PMID: [24669847](https://pubmed.ncbi.nlm.nih.gov/24669847/)
30. Clerissi C, Desdevises Y, Grimsley NH. Prasinoviruses of the marine green alga *Ostreococcus tauri* are mainly species-specific. *J Virol*. 2012; 86: 4611–4619. doi: [10.1128/JVI.07221-11](https://doi.org/10.1128/JVI.07221-11) PMID: [22318150](https://pubmed.ncbi.nlm.nih.gov/22318150/)
31. Winter C, Bouvier T, Weinbauer M, Thingstad T. Trade-Offs between Competition and Defense Specialists among Unicellular Planktonic Organisms: the “Killing the Winner” Hypothesis Revisited. *Microbiology and Molecular Biology Reviews*. 2010; 74: 42–57. doi: [10.1128/MMBR.00034-09](https://doi.org/10.1128/MMBR.00034-09) PMID: [20197498](https://pubmed.ncbi.nlm.nih.gov/20197498/)
32. Bratbak G, Egge JK, Heldal M. Viral mortality of the marine alga *Emiliania huxleyi* (Haptophyceae) and termination of algal blooms. *Mar Ecol-Prog Ser*. 1993; 93: 39–48. doi: [10.3354/meps093039](https://doi.org/10.3354/meps093039)
33. Brussaard CPD, Kuipers B, Veldhuis MJW. A mesocosm study of *Phaeocystis globosa* population dynamics—1. Regulatory role of viruses in bloom. *Harmful Algae*. 2005; 4: 859–874.
34. Hingamp P, Grimsley N, Acinas SG, Clerissi C, Subirana L, Poulain J, et al. Exploring nucleo-cytoplasmic large DNA viruses in Tara Oceans microbial metagenomes. *ISME J*. 2013; 7: 1678–1695. doi: [10.1038/ismej.2013.59](https://doi.org/10.1038/ismej.2013.59) PMID: [23575371](https://pubmed.ncbi.nlm.nih.gov/23575371/)
35. Martinez-Garcia M, Brazel D, Poulton NJ, Swan BK, Gomez ML, Masland D, et al. Unveiling in situ interactions between marine protists and bacteria through single cell sequencing. *The ISME Journal*. 2012; 6: 703–707. doi: [10.1038/ismej.2011.126](https://doi.org/10.1038/ismej.2011.126) PMID: [21938022](https://pubmed.ncbi.nlm.nih.gov/21938022/)
36. Short CM, Rusanova O, Short SM. Quantification of virus genes provides evidence for seed-bank populations of phycodnaviruses in Lake Ontario, Canada. *ISME J*. 2011; 5: 810–821. doi: [10.1038/ismej.2010.183](https://doi.org/10.1038/ismej.2010.183) PMID: [21124493](https://pubmed.ncbi.nlm.nih.gov/21124493/)
37. Thyrhaug R. Stable coexistence in marine algal host-virus systems. *Marine Ecology-Progress Series*. 2003; 254: 27–35.
38. Zingone A, Natale F, Biffali E, Borra M, Forlani G, Sarno D. Diversity in morphology, infectivity, molecular characteristics and induced host resistance between two viruses infecting *Micromonas pusilla*. *Aquatic Microbial Ecology*. 2006; 45: 1–14.
39. Waters RE, Chan AT. *Micromonas pusilla* Virus—the Virus Growth-Cycle and Associated Physiological Events within the Host-Cells—Host Range Mutation. *Journal of General Virology*. 1982; 63: 199–206.
40. Frada M, Probert I, Allen MJ, Wilson WH, de Vargas C. The “Cheshire Cat” escape strategy of the coccolithophore *Emiliania huxleyi* in response to viral infection. *Proc Natl Acad Sci USA*. 2008; 105: 15944–15949. doi: [10.1073/pnas.0807707105](https://doi.org/10.1073/pnas.0807707105) PMID: [18824682](https://pubmed.ncbi.nlm.nih.gov/18824682/)
41. Tomaru Y, Mizumoto H, Nagasaki K. Virus resistance in the toxic bloom-forming dinoflagellate *Heterocapsa circularisquama* to single-stranded RNA virus infection. *Environ Microbiol*. 2009; 11: 2915–2923. doi: [10.1111/j.1462-2920.2009.02047.x](https://doi.org/10.1111/j.1462-2920.2009.02047.x) PMID: [19735281](https://pubmed.ncbi.nlm.nih.gov/19735281/)
42. Simon L, Voisin M, Tatout C, Probst AV. Structure and Function of Centromeric and Pericentromeric Heterochromatin in *Arabidopsis thaliana*. *Front Plant Sci*. 2015; 1049. doi: [10.3389/fpls.2015.01049](https://doi.org/10.3389/fpls.2015.01049) PMID: [26648952](https://pubmed.ncbi.nlm.nih.gov/26648952/)
43. Biscotti MA, Olmo E, Heslop-Harrison JS (Pat). Repetitive DNA in eukaryotic genomes. *Chromosome Res*. 2015; 23: 415–420. doi: [10.1007/s10577-015-9499-z](https://doi.org/10.1007/s10577-015-9499-z) PMID: [26514350](https://pubmed.ncbi.nlm.nih.gov/26514350/)
44. Shibata Y, Yamashita Y, Ozaki K, Nakano Y, Koga T. Expression and Characterization of Streptococcal rgp Genes Required for Rhamnan Synthesis in *Escherichia coli*. *Infect Immun*. 2002; 70: 2891–2898. doi: [10.1128/IAI.70.6.2891-2898.2002](https://doi.org/10.1128/IAI.70.6.2891-2898.2002) PMID: [12010977](https://pubmed.ncbi.nlm.nih.gov/12010977/)
45. Tao J, Feng L, Guo H, Li Y, Wang L. The O-antigen gene cluster of *Shigella boydii* O11 and functional identification of its wzy gene. *FEMS Microbiol Lett*. 2004; 234: 125–132. doi: [10.1111/j.1574-6968.2004.tb09523.x](https://doi.org/10.1111/j.1574-6968.2004.tb09523.x) PMID: [15109730](https://pubmed.ncbi.nlm.nih.gov/15109730/)
46. Feng L, Senchenkova SN, Wang W, Shashkov AS, Liu B, Shevelev SD, et al. Structural and genetic characterization of the *Shigella boydii* type 18 O antigen. *Gene*. 2005; 355: 79–86. doi: [10.1016/j.gene.2005.06.001](https://doi.org/10.1016/j.gene.2005.06.001) PMID: [16055280](https://pubmed.ncbi.nlm.nih.gov/16055280/)

47. Blanc G, Mozar M, Agarkova IV, Gurnon JR, Yanai-Balser G, Rowe JM, et al. Deep RNA Sequencing Reveals Hidden Features and Dynamics of Early Gene Transcription in *Paramecium bursaria* Chlorella Virus 1. PLOS ONE. 2014; 9: e90989. doi: [10.1371/journal.pone.0090989](https://doi.org/10.1371/journal.pone.0090989) PMID: [24608750](https://pubmed.ncbi.nlm.nih.gov/24608750/)
48. Fruscione F, Sturla L, Duncan G, Etten JLV, Valbuzzi P, Flora AD, et al. Differential Role of NADP+ and NADPH in the Activity and Structure of GDP-D-mannose 4,6-Dehydratase from Two Chlorella Viruses. J Biol Chem. 2008; 283: 184–193. doi: [10.1074/jbc.M706614200](https://doi.org/10.1074/jbc.M706614200) PMID: [17974560](https://pubmed.ncbi.nlm.nih.gov/17974560/)
49. Tonetti M, Zanardi D, Gurnon JR, Fruscione F, Armirotti A, Damonte G, et al. *Paramecium bursaria* chlorella virus 1 encodes two enzymes involved in the biosynthesis of GDP-L-fucose and GDP-D-rhamnose. J Biol Chem. 2003; 278: 21559–65. doi: [10.1074/jbc.M301543200](https://doi.org/10.1074/jbc.M301543200) PMID: [12679342](https://pubmed.ncbi.nlm.nih.gov/12679342/)
50. De Castro CD, Molinaro A, Piacente F, Gurnon JR, Sturiale L, Palmigiano A, et al. Structure of N-linked oligosaccharides attached to chlorovirus PBCV-1 major capsid protein reveals unusual class of complex N-glycans. PNAS. 2013; 110: 13956–13960. doi: [10.1073/pnas.1313005110](https://doi.org/10.1073/pnas.1313005110) PMID: [23918378](https://pubmed.ncbi.nlm.nih.gov/23918378/)
51. Wang M, Lai E. Pulsed field separation of large supercoiled and open-circular DNAs and its application to bacterial artificial chromosome cloning. Electrophoresis. 1995; 16: 1–7. doi: [10.1002/elps.1150160102](https://doi.org/10.1002/elps.1150160102) PMID: [7737080](https://pubmed.ncbi.nlm.nih.gov/7737080/)
52. Jancek S, Gourbiere S, Moreau H, Piganeau G. Clues about the Genetic Basis of Adaptation Emerge from Comparing the Proteomes of Two *Ostreococcus* Ecotypes (Chlorophyta, Prasinophyceae). Molecular Biology and Evolution. 2008; 25: 2293–2300. doi: [10.1093/molbev/msn168](https://doi.org/10.1093/molbev/msn168) PMID: [18678753](https://pubmed.ncbi.nlm.nih.gov/18678753/)
53. Cohen M. Notable Aspects of Glycan-Protein Interactions. Biomolecules. 2015; 5: 2056–2072. doi: [10.3390/biom5032056](https://doi.org/10.3390/biom5032056) PMID: [26340640](https://pubmed.ncbi.nlm.nih.gov/26340640/)
54. Denes T, den Bakker HC, Tokman JI, Guldemann C, Wiedmann M. Selection and Characterization of Phage-Resistant Mutant Strains of *Listeria monocytogenes* Reveal Host Genes Linked to Phage Adsorption. Appl Environ Microbiol. 2015; 81: 4295–4305. doi: [10.1128/AEM.00087-15](https://doi.org/10.1128/AEM.00087-15) PMID: [25888172](https://pubmed.ncbi.nlm.nih.gov/25888172/)
55. Eugster MR, Morax LS, Hüls VJ, Huwiler SG, Leclercq A, Lecuit M, et al. Bacteriophage predation promotes serovar diversification in *Listeria monocytogenes*. Molecular Microbiology. 2015; 97: 33–46. doi: [10.1111/mmi.13009](https://doi.org/10.1111/mmi.13009) PMID: [25825127](https://pubmed.ncbi.nlm.nih.gov/25825127/)
56. Lee J-B, Hayashi K, Hayashi T, Sankawa U, Maeda M. Antiviral Activities against HSV-1, HCMV, and HIV-1 of Rhamnan Sulphate from *Monostroma latissimum*. Planta Medica. 1999; 65: 439–441. doi: [10.1055/s-2006-960804](https://doi.org/10.1055/s-2006-960804) PMID: [10418331](https://pubmed.ncbi.nlm.nih.gov/10418331/)
57. Ng PSK, Böhm R, Hartley-Tassell LE, Steen JA, Wang H, Lukowski SW, et al. Ferrets exclusively synthesize Neu5Ac and express naturally humanized influenza A virus receptors. Nat Commun. 2014; 5: 5750. doi: [10.1038/ncomms6750](https://doi.org/10.1038/ncomms6750) PMID: [25517696](https://pubmed.ncbi.nlm.nih.gov/25517696/)
58. Alexander WS, Viney EM, Zhang J-G, Metcalf D, Kauppi M, Hyland CD, et al. Thrombocytopenia and kidney disease in mice with a mutation in the C1galt1 gene. PNAS. 2006; 103: 16442–16447. doi: [10.1073/pnas.0607872103](https://doi.org/10.1073/pnas.0607872103) PMID: [17062753](https://pubmed.ncbi.nlm.nih.gov/17062753/)
59. Farinas B, Mary C, Manes CLD, Bhaud Y, Peaucellier G, Moreau H. Natural synchronisation for the study of cell division in the green unicellular alga *Ostreococcus tauri*. Plant Molecular Biology. 2006; 60: 277–292. doi: [10.1007/s11103-005-4066-1](https://doi.org/10.1007/s11103-005-4066-1) PMID: [16429264](https://pubmed.ncbi.nlm.nih.gov/16429264/)
60. Piacente F, Gaglianone M, Laugieri ME, Tonetti MG. The Autonomous Glycosylation of Large DNA Viruses. International Journal of Molecular Sciences. 2015; 16: 29315–29328. doi: [10.3390/ijms161226169](https://doi.org/10.3390/ijms161226169) PMID: [26690138](https://pubmed.ncbi.nlm.nih.gov/26690138/)
61. Melayah D, Bonnard E, Chalhoub B, Audeon C, Grandbastien M-A. The mobility of the tobacco Tnt1 retrotransposon correlates with its transcriptional activation by fungal factors. Plant J. 2001; 28: 159–168. doi: [10.1046/j.1365-313X.2001.01141.x](https://doi.org/10.1046/j.1365-313X.2001.01141.x) PMID: [11722759](https://pubmed.ncbi.nlm.nih.gov/11722759/)
62. Grandbastien M-A. LTR retrotransposons, handy hitchhikers of plant regulation and stress response. Biochimica et Biophysica Acta (BBA)—Gene Regulatory Mechanisms. 2015; 1849: 403–416. doi: [10.1016/j.bbagr.2014.07.017](https://doi.org/10.1016/j.bbagr.2014.07.017) PMID: [25086340](https://pubmed.ncbi.nlm.nih.gov/25086340/)
63. Miousse IR, Chalbot M-CG, Lumen A, Ferguson A, Kavouras IG, Koturbash I. Response of transposable elements to environmental stressors. Mutation Research/Reviews in Mutation Research. 2015; 765: 19–39. doi: [10.1016/j.mrrev.2015.05.003](https://doi.org/10.1016/j.mrrev.2015.05.003) PMID: [26281766](https://pubmed.ncbi.nlm.nih.gov/26281766/)
64. Iwasaki M, Paszkowski J. Epigenetic memory in plants. The EMBO Journal. 2014; 33: 1987–1998. doi: [10.15252/embj.201488883](https://doi.org/10.15252/embj.201488883) PMID: [25104823](https://pubmed.ncbi.nlm.nih.gov/25104823/)
65. Casacuberta JM, Santiago N. Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene. 2003; 311: 1–11. doi: [10.1016/S0378-1119\(03\)00557-2](https://doi.org/10.1016/S0378-1119(03)00557-2) PMID: [12853133](https://pubmed.ncbi.nlm.nih.gov/12853133/)

66. Lescot M, Hingamp P, Kojima KK, Villar E, Romac S, Veluchamy A, et al. Reverse transcriptase genes are highly abundant and transcriptionally active in marine plankton assemblages. *ISME J.* 2015; doi: [10.1038/ismej.2015.192](https://doi.org/10.1038/ismej.2015.192) PMID: [26613339](https://pubmed.ncbi.nlm.nih.gov/26613339/)
67. Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature.* 2011; 472: 115–119. doi: [10.1038/nature09861](https://doi.org/10.1038/nature09861) PMID: [21399627](https://pubmed.ncbi.nlm.nih.gov/21399627/)
68. Downen RH, Pelizzola M, Schmitz RJ, Lister R, Downen JM, Nery JR, et al. Widespread dynamic DNA methylation in response to biotic stress. *PNAS.* 2012; 109: E2183–E2191. doi: [10.1073/pnas.1209329109](https://doi.org/10.1073/pnas.1209329109) PMID: [22733782](https://pubmed.ncbi.nlm.nih.gov/22733782/)
69. Huff JT, Zilberman D. Dnmt1-Independent CG Methylation Contributes to Nucleosome Positioning in Diverse Eukaryotes. *Cell.* 2014; 156: 1286–1297. doi: [10.1016/j.cell.2014.01.029](https://doi.org/10.1016/j.cell.2014.01.029) PMID: [24630728](https://pubmed.ncbi.nlm.nih.gov/24630728/)
70. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.* 2010; 11: 204–220. doi: [10.1038/nrg2719](https://doi.org/10.1038/nrg2719) PMID: [20142834](https://pubmed.ncbi.nlm.nih.gov/20142834/)
71. Bire S, Casteret S, Piégu B, Beauclair L, Moiré N, Arensbuger P, et al. Mariner Transposons Contain a Silencer: Possible Role of the Polycomb Repressive Complex 2. *PLOS Genet.* 2016; 12: e1005902. doi: [10.1371/journal.pgen.1005902](https://doi.org/10.1371/journal.pgen.1005902) PMID: [26939020](https://pubmed.ncbi.nlm.nih.gov/26939020/)
72. Aasland R, Francis Stewart A, Gibson T. The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional corepressor N-CoR and TFIIB. *Trends Biochem Sci.* 1996; 21: 87–88. doi: [10.1016/S0968-0004\(96\)30009-1](https://doi.org/10.1016/S0968-0004(96)30009-1) PMID: [8882580](https://pubmed.ncbi.nlm.nih.gov/8882580/)
73. Geiduschek EP, Kassavetis GA. The RNA polymerase III transcription apparatus. *J Molec Biol.* 2001; 310: 1–26. doi: [10.1006/jmbi.2001.4732](https://doi.org/10.1006/jmbi.2001.4732) PMID: [11419933](https://pubmed.ncbi.nlm.nih.gov/11419933/)
74. Avrani S, Wurtzel O, Sharon I, Sorek R, Lindell D. Genomic island variability facilitates *Prochlorococcus*–virus coexistence. *Nature.* 2011; 474: 604–608. doi: [10.1038/nature10172](https://doi.org/10.1038/nature10172) PMID: [21720364](https://pubmed.ncbi.nlm.nih.gov/21720364/)
75. Penadés JR, Christie GE. The Phage-Inducible Chromosomal Islands: A Family of Highly Evolved Molecular Parasites. *Annu Rev Virol.* 2015; 2: 181–201. doi: [10.1146/annurev-virology-031413-085446](https://doi.org/10.1146/annurev-virology-031413-085446) PMID: [26958912](https://pubmed.ncbi.nlm.nih.gov/26958912/)
76. Schmidt SM, Panstruga R. Pathogenomics of fungal plant parasites: what have we learnt about pathogenesis? *Curr Opin Plant Biol.* 2011; 14: 392–399. doi: [10.1016/j.pbi.2011.03.006](https://doi.org/10.1016/j.pbi.2011.03.006) PMID: [21458359](https://pubmed.ncbi.nlm.nih.gov/21458359/)
77. Han Y, Liu X, Benny U, Kistler HC, VanEtten HD. Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. *The Plant Journal.* 2001; 25: 305–314. doi: [10.1046/j.1365-3113x.2001.00969.x](https://doi.org/10.1046/j.1365-3113x.2001.00969.x) PMID: [11208022](https://pubmed.ncbi.nlm.nih.gov/11208022/)
78. Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, et al. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature.* 2010; 464: 367–373. doi: [10.1038/nature08850](https://doi.org/10.1038/nature08850) PMID: [20237561](https://pubmed.ncbi.nlm.nih.gov/20237561/)
79. Holoch D, Moazed D. RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet.* 2015; 16: 71–84. doi: [10.1038/nrg3863](https://doi.org/10.1038/nrg3863) PMID: [25554358](https://pubmed.ncbi.nlm.nih.gov/25554358/)
80. Bond DM, Baulcombe DC. Small RNAs and heritable epigenetic variation in plants. *Trends Cell Biol.* 2014; 24: 100–107. doi: [10.1016/j.tcb.2013.08.001](https://doi.org/10.1016/j.tcb.2013.08.001) PMID: [24012194](https://pubmed.ncbi.nlm.nih.gov/24012194/)
81. Cerutti H, Ma X, Msanne J, Repas T. RNA-Mediated Silencing in Algae: Biological Roles and Tools for Analysis of Gene Function. *Eukaryotic Cell.* 2011; 10: 1164–1172. doi: [10.1128/EC.05106-11](https://doi.org/10.1128/EC.05106-11) PMID: [21803865](https://pubmed.ncbi.nlm.nih.gov/21803865/)
82. Giovannoni SJ, Cameron Thrash J, Temperton B. Implications of streamlining theory for microbial ecology. *ISME J.* 2014; 8: 1553–1565. doi: [10.1038/ismej.2014.60](https://doi.org/10.1038/ismej.2014.60) PMID: [24739623](https://pubmed.ncbi.nlm.nih.gov/24739623/)
83. Billmyre RB, Calo S, Feretzaki M, Wang X, Heitman J. RNAi function, diversity, and loss in the fungal kingdom. *Chromosome Res.* 2013; 21: 561–572. doi: [10.1007/s10577-013-9388-2](https://doi.org/10.1007/s10577-013-9388-2) PMID: [24173579](https://pubmed.ncbi.nlm.nih.gov/24173579/)
84. Clerissi C, Grimsley N, Subirana L, Maria E, Oriol L, Ogata H, et al. Prasinovirus distribution in the Northwest Mediterranean Sea is affected by the environment and particularly by phosphate availability. *Virology.* 2014; 466–467: 146–157. doi: [10.1016/j.virol.2014.07.016](https://doi.org/10.1016/j.virol.2014.07.016) PMID: [25109909](https://pubmed.ncbi.nlm.nih.gov/25109909/)
85. Baudoux A-C, Lebretonchel H, Dehmer H, Latimier M, Ederm R, Rigaut-Jalabert F, et al. Interplay between the genetic clades of *Micromonas* and their viruses in the Western English Channel. *Environ Microbiol Rep.* 2015; 7: 765–773. doi: [10.1111/1758-2229.12309](https://doi.org/10.1111/1758-2229.12309) PMID: [26081716](https://pubmed.ncbi.nlm.nih.gov/26081716/)
86. Vault D, Le Gall F, Marie D, Guillou L, Partensky F. The Roscoff Culture Collection (RCC): a collection dedicated to marine picoplankton. *Nova Hedwigia.* 2004; 79: 49–70.
87. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 2012; 28: 1647–1649. doi: [10.1093/bioinformatics/bts199](https://doi.org/10.1093/bioinformatics/bts199) PMID: [22543367](https://pubmed.ncbi.nlm.nih.gov/22543367/)
88. Geneious—Home page [Internet]. [cited 10 May 2016]. <http://www.geneious.com/>

89. Grimsley N, Pequin B, Bachy C, Moreau H, Piganeau G. Cryptic Sex in the Smallest Eukaryotic Marine Green Alga. *Molec Biol Evol*. 2010; 27: 47–54. doi: [10.1093/molbev/msp203](https://doi.org/10.1093/molbev/msp203) PMID: [19734297](https://pubmed.ncbi.nlm.nih.gov/19734297/)
90. Babraham Bioinformatics—FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. [cited 10 May 2016]. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
91. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*. 2013; 14: R36. doi: [10.1186/gb-2013-14-4-r36](https://doi.org/10.1186/gb-2013-14-4-r36) PMID: [23618408](https://pubmed.ncbi.nlm.nih.gov/23618408/)
92. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015; 31: 166–169. doi: [10.1093/bioinformatics/btu638](https://doi.org/10.1093/bioinformatics/btu638) PMID: [25260700](https://pubmed.ncbi.nlm.nih.gov/25260700/)
93. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15: 1–21. doi: [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8) PMID: [25516281](https://pubmed.ncbi.nlm.nih.gov/25516281/)
94. Rodriguez F, Derelle E, Guillou L, Le Gall F, Vaulot D, Moreau H. Ecotype diversity in the marine picoeukaryote *Ostreococcus* (Chlorophyta, Prasinophyceae). *Environ Microbiol*. 2005; 7: 853–859. doi: [10.1111/j.1462-2920.2005.00758.x](https://doi.org/10.1111/j.1462-2920.2005.00758.x) PMID: [15892704](https://pubmed.ncbi.nlm.nih.gov/15892704/)
95. ORCAE: Online Resource for Community Annotation of Eukaryotes [Internet]. [cited 10 May 2016]. <http://bioinformatics.psb.ugent.be/orcae/overview/OsttaV2>
96. Sterck L, Billiau K, Abeel T, Rouz  P, Van de Peer Y. ORCAE: online resource for community annotation of eukaryotes. *Nat Meth*. 2012; 9: 1041–1041. doi: [10.1038/nmeth.2242](https://doi.org/10.1038/nmeth.2242) PMID: [23132114](https://pubmed.ncbi.nlm.nih.gov/23132114/)