Emergent RNA–RNA interactions can promote stability in a facultative phototrophic endosymbiosis

Benjamin H. Jenkins[a,b,1], Finlay Maguire[c], Guy Leonard[a,b], Joshua D. Eaton[a], Steven West[a], Benjamin E. Housden[d], David S. Milner[a,b], and Thomas A. Richards[a,b,1]

[a]Living Systems Institute and Biosciences, University of Exeter, Devon EX4 4QD, United Kingdom; [b]Department of Zoology, University of Oxford, Oxford OX1 3SZ, United Kingdom; [c]Faculty of Computer Science, Dalhousie University, Halifax NS B3H 1W5, Canada; and [d]Living Systems Institute and College of Medicine and Health, University of Exeter, Devon EX4 4QD, United Kingdom

Eukaryote–eukaryote endosymbiosis was responsible for the spread of chloroplast (plastid) organelles. Stability is required for the metabolic and genetic integration that drives the establishment of new organelles, yet the mechanisms that act to stabilize emergent endosymbioses—between two fundamentally selfish biological organisms—are unclear. Theory suggests that enforcement mechanisms, which punish misbehavior, may act to stabilize such interactions by resolving conflict. However, how such mechanisms can emerge in a facultative endosymbiosis has yet to be explored. Here, we propose that endosymbiont–host RNA–RNA interactions, arising from digestion of the endosymbiont population, can result in a cost to host growth for breakdown of the endosymbiosis. Using the model facultative endosymbiosis between *Paramecium bursaria* and *Chlorella* spp., we demonstrate that this mechanism is dependent on the host RNA-interference (RNAi) system. We reveal through small RNA (sRNA) sequencing that endosymbiont-derived messenger RNA (mRNA) released upon endosymbiont digestion can be processed by the host RNAi system into 23-nt sRNA. We predict multiple regions of shared sequence identity between endosymbiont and host mRNA, and demonstrate through delivery of synthetic endosymbiont sRNA that exposure to these regions can knock down expression of complementary host genes, resulting in a cost to host growth. This process of host gene knockdown in response to endosymbiont-derived RNA processing by host RNAi factors, which we term "RNAi collisions," represents a mechanism that can promote stability in a facultative eukaryote–eukaryote endosymbiosis. Specifically, by imposing a cost for breakdown of the endosymbiosis, endosymbiont–host RNA–RNA interactions may drive maintenance of the symbiosis across fluctuating ecological conditions.

coevolution | symbiosis | protist | algae | cell–cell interaction

A nascent endosymbiosis between eukaryotes was a prerequisite for the spread of photosynthetic organelles, such as plastids (1–5). This transition from transiently engulfed cells to obligate organelles is driven by metabolic and genetic integration. Yet, to become manifest, a stable intermediary state must exist upon which this process of integration can proceed (1, 3, 5–8). Conflict is an inevitable outcome of all symbioses, the resolution of which can significantly impact the stability of an interaction. Enforcement mechanisms that punish misbehavior can act to stabilize symbioses (9, 10), yet we know little about how these can arise in an emergent endosymbiotic interaction. *Paramecium bursaria*—a ciliate protist that harbors a population of often clonal intracellular green algae, *Chlorella* spp. (11–13)—represents a tractable model system to study emergent mechanisms in endosymbiosis (14). The interaction is facultative (11, 14–17) and based on two-way metabolic exchange (18–26). Endosymbiotic algae are housed within modified host phagosomes called perialgal vacuoles, which may be fused with host lysosomes to trigger digestion (14, 27), allowing *P. bursaria* to maintain a high degree of control over the interaction in the event of conflict (28–30). However, it is unclear how this

endosymbiotic system is protected from overexploitation by the host, which would ultimately lead the interaction to collapse (31–36).

RNA–RNA interactions can play a role in host–pathogen symbiotic systems (37–41), whereby RNA can "hijack" the small interfering RNA (siRNA) pathway of the symbiotic partner to modulate expression of genes involved in virulence or alternatively resistance. Whether analogous RNA–RNA interactions could occur in an endosymbiotic system has yet to be elucidated. The presence of a functional siRNA pathway in *Paramecium*, and its role in RNA interference (RNAi), has been validated as a tool for gene silencing in both *P. bursaria* (42) and the non-photoendosymbiotic congener *Paramecium tetraurelia* (43–47). RNAi can be initiated through the provision of bacterial food transformed to express double-stranded RNA (dsRNA) with high sequence similarity to a target transcript (42, 44). *Paramecium* can also process single-stranded RNA (ssRNA)—including ribosomal RNA (rRNA) and messenger RNA (mRNA)—derived from a prokaryotic cell acquired through phagotrophy (43, 48). This processing is mediated by conserved RNAi protein components, many of which also function in endogenous transcriptome regulation (42, 43, 45, 48, 49). Significantly, these studies propose

Significance

Stable endosymbiosis between eukaryotic microbes has driven the evolution of further cellular complexity. Yet the mechanisms that can act to stabilize an emergent eukaryote–eukaryote endosymbiosis are unclear. Using the model facultative endosymbiotic system, *Paramecium bursaria*, we demonstrate that endosymbiont–host RNA–RNA interactions can drive a cost to host growth upon endosymbiont digestion. These RNA–RNA interactions are facilitated by the host RNA-interference system. For endosymbiont messenger RNA sharing a high level of sequence identity with host transcripts, this process can result in host gene knockdown. We propose that these endosymbiont–host RNA–RNA interactions—"RNA-interference collisions"—represent an emergent mechanism to sanction the host for breakdown of the endosymbiosis, promoting the stability of the facultative endosymbiotic interaction.

EVOLUTION

that mRNA, derived both exogenously (42, 43) and endogenously (42, 45), can act as substrates for siRNA generation in *Paramecium*.

Here, we demonstrate that RNA released upon digestion of the algal endosymbiont is processed by the host RNAi system in *P. bursaria*. For endosymbiont-derived mRNA sharing a high level of sequence identity with host transcripts, this processing interferes with endogenous host gene expression, resulting in a cost to host growth (*SI Appendix*, Fig. S1). We track the interaction through small RNA (sRNA) sequencing, recapitulate the effect through exposure to synthetic endosymbiont RNA, and demonstrate that this mechanism is mediated by host Dicer, AGO-Piwi, Pds1, and RdRP proteins. This process of host gene knockdown in response to endosymbiont-derived RNA processing by host RNAi factors, which we term "RNAi collisions," represents an emergent mechanism which can promote stability in a facultative eukaryote–eukaryote endosymbiosis. By imposing a cost for breakdown of the endosymbiosis, endosymbiont–host RNA–RNA interactions may drive maintenance of the symbiosis across fluctuating ecological conditions and symbiotic status.

Endosymbiont Digestion in *P. bursaria* Results in an RNAi-Mediated "Physiological Cost" to the Host.
P. bursaria can be purged of endosymbiotic algae via treatment with the ribosomal translational inhibitor, cycloheximide (14). A comparison of ribosomal protein (RP) L29A amino acid position, the active site of cycloheximide function, confirmed that *Paramecium* possess a specific nucleotide polymorphism identified as a determinant of cycloheximide resistance in other species (*SI Appendix*, Fig. S2). This substitution was absent in all green algal species assessed, including known algal endosymbionts of *P. bursaria*. Upon treatment with cycloheximide, a significant reduction in algal-chlorophyll fluorescent intensity per host cell was observed after 2 to 3 d (Fig. 1A). A clear decoupling of host *P. bursaria* cell number and algal fluorescence was also observed, consistent with translational inhibition in the algae but not in the host (*SI Appendix*, Fig. S2), demonstrating that loss of a photosynthetic endosymbiont population does not immediately result in a decline in host cell number. Additional staining with LysoTracker Green to identify acidic vacuoles, including lysosomes, revealed that this loss of algal fluorescence was linked to increased lysosomal activity in the host cytoplasmic environment (Fig. 1B). These data support that elimination of the endosymbiotic algal population during cycloheximide treatment (14, 27) is triggered by host digestion.

Continued treatment with cycloheximide resulted in a significant retardation to *P. bursaria* culture growth; however, this same effect was not observed in the nonphotoendosymbiotic congener species, *P. tetraurelia* (Fig. 1C). Both *Paramecium* species harbor the same cycloheximide resistance-conferring substitution in RPL29A (*SI Appendix*, Fig. S2), suggesting that elimination of endosymbiotic algae is costly to *P. bursaria* culture growth. Loss of algal-derived photosynthate and other plastid-derived metabolites represents an obvious cost. However, we sought to explore whether a part of this cost may be attributed to host exposure to endosymbiont-derived RNA during digestion of the endosymbiont population, and whether this cost was mediated by the host RNAi system.

Knockdown of *Dcr1* (a host-encoded endoribonuclease Dicer required for siRNA generation) through complementary dsRNA exposure significantly rescued the cost to *P. bursaria* culture growth associated with cycloheximide treatment (Fig. 1D and *SI Appendix*, Fig. S3). This was consistent with *Dcr1* (Dicer) knockdown in a prior study attenuating the effect of *Escherichia coli* vector–based RNAi feeding in *P. bursaria* (42). In the aforementioned study, knockdown of host-encoded *PiwiA1*, *PiwiC1* (AGO-Piwi effectors required for targeted RNA cleavage), and *Pds1* (a *Paramecium*-specific component with an unknown but essential role in exogenously induced RNAi) also attenuated an

E. coli–vector feeding–based RNAi effect (42). Here, knockdown of host-encoded *PiwiA1* and *Pds1* similarly rescued the cost to *P. bursaria* culture growth associated with cycloheximide treatment (Fig. 1E and *SI Appendix*, Fig. S3). Involvement of *Pds1* confirms that the observed RNAi effect is mediated by the host, as no identifiable homolog of *Pds1* could be identified in the green algal genomes and transcriptomes sampled (42, 50). Furthermore, it excludes the possibility that off-target effects arising from host Dicer knockdown, including potential compensatory function of additional Dicer or Dicer-like paralogues in *P. bursaria* (42), could be responsible for the RNAi effect observed. *Pds1* has previously been inferred to function in import of RNA across the digestive vacuole of *P. tetraurelia* (43), and thus our work here and elsewhere (42) represents a step closer to elucidating the function of this protein in an endosymbiotic system. This is an important factor for understanding how systems have evolved to control transmission of RNA across membranes between interacting organisms (37, 39, 43, 51).

It is important to note that recursive knockdown of Dicer, *PiwiA1*, and *Pds1* through RNAi represents a potential paradox and so constitutes only partial loss of function (42, 52). As such, it is unclear whether the partial rescue to culture growth observed is due to incomplete perturbation of RNAi function, disruption of favorable RNAi functions elsewhere in the cellular system, or additional factors, such as loss of algal-derived metabolites. However, the near-complete rescue arising from partial knockdown of Dicer suggests that it is host Dicer processing, not loss of algal-derived metabolites, that represents the more significant factor relating to growth retardation under these conditions. Taken together, these data suggest that the physiological cost to *P. bursaria* growth incurred during cycloheximide treatment, in which the endosymbiotic algae are being broken down and digested by the host, is mediated by host-encoded RNAi components. These data support the occurrence of RNAi-mediated RNA–RNA interactions between endosymbiont and host.

Endosymbiont Breakdown Triggers an Abundance of Dicer-Dependent Endosymbiont-Derived sRNA within *P. bursaria*.
To investigate the occurrence of RNA–RNA interactions between endosymbiont and host, we tracked the abundance of endosymbiont-derived, host-processed sRNA in *P. bursaria* during endosymbiont digestion. Disruption of host RNAi was achieved through recursive knockdown of Dicer, allowing us to directly test for an increase in endosymbiotic algal-derived 23-nt sRNA [the size associated with host Dicer processing (42, 43, 53)] resulting from cycloheximide induced endosymbiont breakdown. Dicer was chosen for further work to assess this effect, as knockdown of this component resulted in the greatest rescue to *P. bursaria* culture growth during endosymbiont digestion (Fig. 1D). Upon treatment with cycloheximide, we identified an increased abundance in all 21- to 29-nt reads mapping to endosymbiont-derived mRNA over 2 to 3 d (Fig. 2A and *SI Appendix*, Fig. S4). This same trend was observed for reads mapping to algal endosymbiont rRNA-derived sRNA (*SI Appendix*, Fig. S5). Reads were mapped with 100% identity, allowing no mismatches. Any reads that additionally mapped to the host with 100% identity were removed to ensure that the subset of sRNA detected was of definitive algal origin. A significant increase in algal mRNA-derived 23-nt sense and antisense sRNA demonstrates a greater abundance of potential RNAi substrates during endosymbiont digestion (42, 43, 53) (Fig. 2A).

Partial knockdown of host Dicer during cycloheximide treatment significantly ablated this endosymbiont mRNA-derived 23-nt antisense abundance after 3 d of cycloheximide treatment (Fig. 2B and *SI Appendix*, Fig. S4). This suggests that an increase in endosymbiont mRNA-derived 23-nt antisense sRNA during endosymbiont digestion is dependent on host Dicer function. This is consistent with knockdown of host Dicer specifically reducing 23-nt sRNA abundance in *Paramecium* (42, 43, 53) and the observation
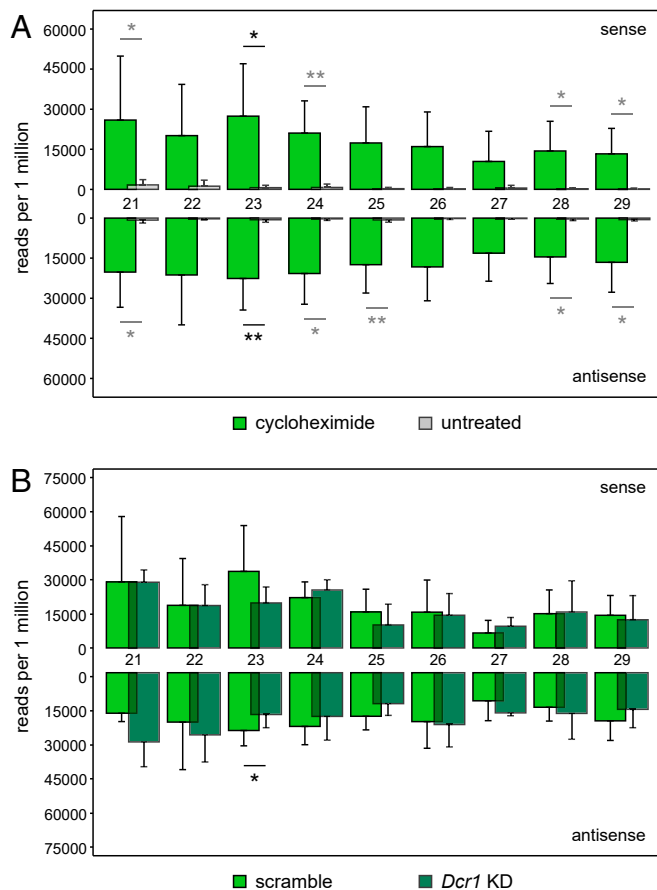
2 of 11 | PNAS
https://doi.org/10.1073/pnas.2108874118

Jenkins et al.
Emergent RNA–RNA interactions can promote stability in a facultative phototrophic endosymbiosis

Fig. 1. Endosymbiont digestion in *P. bursaria* results in an RNAi-mediated "physiological cost" to the host. (*A*) *P. bursaria* cell number (gray) versus algal chlorophyll fluorescent intensity (pink) in stationary phase cultures treated with cycloheximide (50 μg · mL^{-1}) compared to untreated controls. Loss of algal fluorescent intensity indicates algal death/loss of function in response to cycloheximide treatment. Data are represented as mean ± SD of six biological replicates. (*B*) Stages of endosymbiont elimination in representative *P. bursaria* cells (1–6) over 3 d of cycloheximide treatment (50 μg · mL^{-1}). LysoTracker Green fluorescence (green; ') indicates increased host lysosomal activity in response to cycloheximide treatment. Algal chlorophyll fluorescence (pink; ") highlights endosymbiotic algae within the *P. bursaria* cell. Imaging was performed at 10× magnification on an ImageXpress Pico Automated Imaging System to generate composite "tiled" images (see *Methods*). (Scale bar, 50 μm.) (*C*) *Paramecium* cell number in *P. bursaria* or *P. tetraurelia* cultures after 8 d of cycloheximide treatment (50 μg · mL^{-1}; green) compared to untreated controls (light gray). *Paramecium* cultures were concurrently fed with *E. coli* transformed with an empty RNAi vector for 12 d, starting 4 d prior to cycloheximide treatment. (*D*) Percentage change in *P. bursaria* cell number in cultures treated with cycloheximide for 8 d (50 μg · mL^{-1}). *Paramecium* cultures were concurrently fed with *E. coli* expressing Dicer (*Dcr1*) dsRNA (dark gray) to induce knockdown (KD), nonhit "scramble" dsRNA, or an empty vector control (green) for 12 d, starting 4 d prior to cycloheximide treatment. The relative effect of Dicer dsRNA exposure indicates partial rescue of *P. bursaria* culture growth retardation in response to cycloheximide-induced endosymbiont digestion. (*E*) Percentage change in *P. bursaria* cell number in cultures treated with cycloheximide for 6 d (50 μg · mL^{-1}) compared to untreated controls. *Paramecium* cultures were concurrently fed with *E. coli* expressing *PiwiA1*, *PiwiC1*, or *Pds1* dsRNA (dark gray) to induce KD or a nonhit "scramble" dsRNA control (green) for 18 d, starting 12 d prior to cycloheximide treatment. Note that this increased duration of dsRNA feeding prior to cycloheximide treatment is in accordance with a prior study (42), which demonstrated a delayed onset of *PiwiA1*, *PiwiC1*, and *Pds1* KD compared to *Dcr1* KD. (*C–E*) Boxplot data are represented as maximum, upper quartile (Q3), mean, lower quartile (Q1) and minimum values of six biological replicates. Individual data points are shown. Significance was calculated as *$P \le 0.05$, **$P \le 0.01$, and ***$P \le 0.001$ using a generalized linear model with quasi-Poisson distribution. See also *SI Appendix*, Fig. S3 for the raw count data used to calculate the percentage change in cell number presented in *D* and *E*.

Jenkins et al.
Emergent RNA–RNA interactions can promote stability in a facultative phototrophic endosymbiosis

EVOLUTION

Fig. 2. Endosymbiont breakdown triggers an influx of Dicer-dependent endosymbiont-derived sRNA within *P. bursaria*. (*A*) Size distribution (nt) of sRNA mapped to endosymbiont-derived cytoplasmic mRNA. sRNA was extracted from *P. bursaria* cultures over 2 d (day 2 and 3) of cycloheximide treatment (50 μg · mL⁻¹) or from untreated controls. Note the relative increase in 23-nt abundance between untreated and cycloheximide-treated cultures. *P. bursaria* cultures were fed with *E. coli* transformed to express nonhit "scramble" dsRNA. Data are represented as mean ± SD of six biological replicates and normalized against total endosymbiont mRNA mapping 21- to 29-nt reads per dataset. (*B*) Size distribution (nt) of sRNA mapped to endosymbiont-derived cytoplasmic mRNA. sRNA was extracted from *P. bursaria* cultures on day 3 of cycloheximide treatment (50 μg · mL⁻¹). *P. bursaria* cultures were fed with *E. coli* transformed to express Dicer (*Dcr1*) dsRNA to induce knockdown (KD) or a nonhit "scramble" dsRNA control. Note the relative increase in 23-nt abundance during cycloheximide treatment in *P. bursaria* cultures exposed to scramble dsRNA compared to Dicer dsRNA. Data are represented as mean ± SD of three biological replicates and normalized against total endosymbiont mRNA mapping 21- to 29-nt reads per dataset. (*A* and *B*) Feeding was conducted daily for 4 d prior to cycloheximide treatment and continued throughout. Significance was calculated as *$P \leq 0.05$, **$P \leq 0.01$, and ***$P \leq 0.001$ using a generalized linear model with quasi-Poisson distribution. All curated "endosymbiont" mRNA transcript bins used for sRNA mapping are available on Figshare (DOI: 10.6084/m9.figshare.12301736) (77). All reads in this analysis were mapped to the endosymbiont with 100% identity, and any reads that additionally mapped to the host with 100% identity were removed. See also *SI Appendix*, Fig. S4 for sRNA abundance at each individual day.

that *Paramecium* RNAi factors are capable of processing both endogenously and exogenously derived mRNA (43, 45). It is important to note that this effect likely underrepresents the full extent of host Dicer processing, as recursive knockdown of Dicer is paradoxical and so constitutes only partial loss of function (42, 52). Furthermore, these data represent only a subset of the sRNA potentially present due to the stringent mapping approach used to identify endosymbiont-derived sRNA, as any reads that additionally

mapped to a host RNA template with 100% sequence identity in either orientation were excluded from this analysis. While these reads with high shared sequence identity are the most important subset of sRNA for the identification of putative mRNA–mRNA interactions between endosymbiont and host, their exclusion here has allowed us to identify host RNAi processing of definitively endosymbiont-derived transcripts. Interestingly, one of the host-derived transcripts that mapped with 100% identity to an excluded endosymbiont-derived 23-nt sRNA was *P. bursaria* heat shock protein 90 (*HSP90*). An investigation of regions of high shared sequence identity between endosymbiont and host is explored below.

Next, we conducted a series of control observations. Firstly, we assessed whether cycloheximide treatment was altering general host-derived sRNA production, despite the inferred host resistance discussed (*SI Appendix*, Fig. S2). Reads were mapped to a dataset of 20 host transcripts that contained no potential 23-nt overlap with any identified algal transcripts (allowing for ≤2-nt mismatches) to ensure that these host transcripts were unaffected by an increased rate of putative RNA–RNA interactions derived from the endosymbiont (*SI Appendix*, Fig. S6). No significant increase in 23-nt abundance was observed for reads mapping to this subset of host mRNA transcripts during cycloheximide treatment, consistent with host resistance to cycloheximide not altering the host-derived population of sRNAs with low sequence identity to algal mRNA. Secondly, we assessed whether endosymbiotic algal strains cultured under free-living conditions would also generate an abundance of 23-nt sRNA upon cycloheximide treatment (*SI Appendix*, Fig. S7). Importantly, no clear increase in 21- to 25-nt algal-derived sRNA abundance was observed during algal treatment with cycloheximide when grown outside of the host cytoplasmic environment, nor was a 23-nt peak in algal-derived sRNAs evident. This is consistent with data from a previous study which found that the algal endosymbiont of *P. bursaria* was not actively generating sRNA >20-nt (42). These results suggest that increased algal endosymbiont-derived 23-nt sRNA abundance upon treatment with cycloheximide is dependent on host sRNA processing within the endosymbiotic system. Taken together, these data indicate that an abundance of endosymbiont-derived RNAs are released during endosymbiont digestion, which then act as substrates for processing by the host RNAi system, resulting in 23-nt sRNAs. Importantly, processing of endosymbiont-derived sRNA in a host Dicer-dependent manner indicates that these interactions are occurring in the host cytoplasm, supporting the hypothesis that endosymbiont–host mRNA–mRNA interactions are possible.

Comparison of Transcriptome Data Reveals the Potential for Host Transcript Interaction by Endosymbiont-Derived RNAs. To understand the extent of possible RNA–RNA interactions between endosymbiont and host, we built a bioinformatic transcriptome processing tool, eDicer (54), which allows identification of all possible Dicer-generated sense and antisense oligonucleotides produced from a given transcriptome dataset. By mapping 23-nt reads to a second ("host") dataset allowing for ≤2-nt mismatches, potential RNA–RNA interactions between an input (endosymbiont, vector, or food) and a host RNA population can be identified. Inclusion of reads with ≤2-nt mismatches were based on the tolerance for mismatching complementarity reported during RNAi-mediated knockdown of gene expression in multiple systems (55–59). Confirmation of a similar mismatch tolerance in *P. bursaria* is shown in Fig. 3*E* and *SI Appendix*, Fig. S12. For a full overview of the eDicer comparative analysis process, please refer to the *SI Appendix*.

Using a dataset consisting of transcripts binned as either "endosymbiont" or "host" [using a curated *P. bursaria* transcriptome (60)], we identified 35,703 distinct 23-nt putative mRNA–mRNA interactions between the *P. bursaria* "host" and

"endosymbiont" RNA populations, representing 0.121% of the total inventory of distinct host 23-nt k-mers identified (*SI Appendix*, Fig. S8*A*; see also Dataset S2). This was found to be 120-fold greater than the number of putative 23-nt mRNA–mRNA interactions predicted between bacterial food and host transcripts (from two different bacterial sources). Furthermore, the ratio of total:"potentially lethal" putative 23-nt mRNA–mRNA interactions [determined by cross-referencing genes known to be conditionally essential in *Saccharomyces cerevisiae* (61)] was found to be greater between endosymbiont-and-host transcripts (1:1.4) than between bacterial food-and-host transcripts (1:0.16/0.17). Similar patterns were observed for predicted rRNA–rRNA interactions (*SI Appendix*, Fig. S8*B*; see also Dataset S2). These in silico analyses demonstrate that there is far greater potential for the occurrence of both mRNA–mRNA and rRNA–rRNA interactions between the algal endosymbiont and host (eukaryote–eukaryote) RNA populations than there is between bacterial food and host (prokaryote–eukaryote) RNA populations.

In Vivo Exposure to a Synthetic Endosymbiont mRNA-Derived Chimera Generates Knockdown of High-Identity Host Transcripts. Having identified the occurrence of putative endosymbiont–host RNA–RNA interactions in silico we assessed whether exposure to synthetic fragments of algal mRNA could recapitulate the cost to host growth observed during endosymbiont digestion (Fig. 1*E*). We identified 10 endosymbiont mRNA interaction fragments that shared >91% (or 21-nt) sequence identity with host transcripts across a 23-nt region selected at random from the analysis reported in *SI Appendix*, Fig. S8*A*. Each mRNA interaction fragment was chosen to contain at least one single nucleotide polymorphism (SNP) specific to the endosymbiont to ensure that any identifiable effect could be attributed to an endosymbiont-like transcript rather than the host version. These interaction fragments were predicted to still be effective RNAi templates for *P. bursaria* based on the tolerance for mismatching complementarity observed in RNAi-mediated knockdown in other systems (55–59) (and confirmed below in *P. bursaria*; Fig. 3*E* and *SI Appendix*, Fig. S12). Six mRNA interaction fragments showed putative homology to "nonlethal" yeast genes (including *EF1-α* and *HSP90*); two showed putative homology to "lethal" yeast genes (including tubulin-β [*tub-β*]); and the remaining two had no identifiable homologs in yeast (Dataset S1). Significantly, *HSP90* was the host transcript previously identified as a candidate for putative endosymbiont–host RNA–RNA interaction in the sRNA analysis of Fig. 2. However, to ensure that any identifiable effect could be attributed to the endosymbiont-derived transcript, the newly identified interaction fragment was chosen from a different region of this gene containing two SNPs specific to the endosymbiont. All 10 mRNA interaction fragments were composed of the predicted interacting 23-nt sRNA sequence (*SI Appendix*, Fig. S8*A*) flanked by 11-nt of contiguous "nonhit" endosymbiont transcript as a filler on each side (Dataset S3). These ten 45-nt fragments were combined into a single 450-nt synthetic endosymbiotic algal chimera (Fig. 3*A*), cloned into an L4440 plasmid, and transformed into *E. coli* for feeding-induced RNAi.

Exposure to endosymbiotic algal chimera dsRNA resulted in significant retardation of *P. bursaria* culture growth (Fig. 3*B*). Once more, this effect was attenuated by recursive knockdown of host Dicer, demonstrating an RNAi-mediated response to synthetic endosymbiont–host mRNA–mRNA interactions that resembled the host response to endosymbiont digestion in Fig. 1 *C*–*E*. It is important to note that synthetic exposure to endosymbiont-derived RNA via an *E. coli* feeding vector would likely overrepresent these putative RNA interactions. However, to address this issue, we designed a nonhit "nonsense" control composed of a tandem assembly of the 11-nt regions of contiguous "nonhit" algal transcript present in the chimera. A relative dilution of endosymbiont

chimera dsRNA delivery alongside the "nonsense" control (1, 1:1, and 1:3) also resulted in *P. bursaria* culture growth retardation (*SI Appendix*, Fig. S9), demonstrating that a reduced relative delivery of synthetic endosymbiont RNA can also result in a cost to host growth. Similar experiments were conducted to investigate the possibility of endosymbiont–host rRNA–rRNA interactions but showed that there was no equivalent effect on host growth, suggesting that rRNA is shielded from such effects (*SI Appendix*, Fig. S10; see also Dataset S4).

To identify which putative mRNA–mRNA interactions were capable of facilitating a cost to *P. bursaria* culture growth, we identified the ten 450-nt host transcripts predicted to be hit by each endosymbiont-derived 23-nt mRNA interaction fragment used in the chimera (Fig. 3*C*; three examples are shown). Exposure to dsRNA corresponding to each of the individual longer form *P. bursaria* transcripts allowed us to identify a subset of transcript targets which resulted in a significant cost to host growth relative to Dicer knockdown controls (Fig. 3*D*; see *SI Appendix*, Fig. S11 for all 10). These *P. bursaria* transcripts correspond to elongation factor-1α (*EF-1α*), *HSP90* (the host transcript identified as a candidate for putative endosymbiont–host RNA–RNA interaction in the sRNA analysis of Fig. 2), and *tub-β* chain. It can therefore be inferred that these are the 23-nt mRNA interaction fragments (Fig. 3 *A* and *C*) that likely resulted in a cost to host growth observed during synthetic endosymbiont chimera dsRNA exposure (Fig. 3*B*). Using mRNA extracted from *P. bursaria* during the chimera-RNAi feeding described (Fig. 3*B*), qPCR revealed a reduction in host transcript expression of *EF-1α* and *tub-β* in response to endosymbiont chimera dsRNA exposure (Fig. 3*E*). Host transcript expression of *HSP90* appears inconclusive; however, expression of all three host transcripts was partially rescued upon recursive knockdown of host Dicer (*SI Appendix*, Fig. S12). From these data, we can therefore conclude that 23-nt of >91% complementary "endosymbiont" transcript is sufficient to facilitate detectable knockdown of a corresponding host transcript, demonstrating that exposure to endosymbiont-derived RNA originating from phagosomes is capable of impacting host gene expression via host RNAi knockdown.

It is important to consider that delivery of a synthetic endosymbiont-derived RNA chimera represents only an approximation of the putative 23-nt RNA–RNA interactions occurring in *P. bursaria*. Exposure to endosymbiont RNA via an *E. coli* feeding vector would likely overrepresent these putative interactions. However, the random generation of 23-nt Dicer substrates across the dsRNA chimera (only three regions of which resulted in a detectable cost to host growth; Fig. 3 *A* and *D* and *SI Appendix*, Fig. S11) represents only part of the wider *E. coli*–derived RNA population processed by the host RNAi system during chimera-RNAi feeding. Together with the dilution of endosymbiont chimera dsRNA delivery (*SI Appendix*, Fig. S9), these observations support the hypothesis that only a relatively small number of 23-nt RNA–RNA interactions can induce a cost to host growth in *P. bursaria*. Furthermore, these data confirm that sRNAs with ≤2-nt mismatches are effective templates for RNAi-mediated knockdown of gene expression in *P. bursaria*, suggesting that the sRNA analysis in Fig. 2 may underrepresent the true scale of endosymbiont-derived sRNAs that can result in RNA–RNA interactions between endosymbiont and host. Importantly, we note that the occurrence of putative endosymbiont–host RNA–RNA interactions does not exclude a wider cost to host growth arising from processing a large population of algal-derived RNAs with no defined host target, a process that must pose a cost to host cellular economics and transcriptional control/fidelity. For further justification of this "synthetic" approach, and a consideration of the results that we can reliably draw from these data, see the *SI Appendix*.
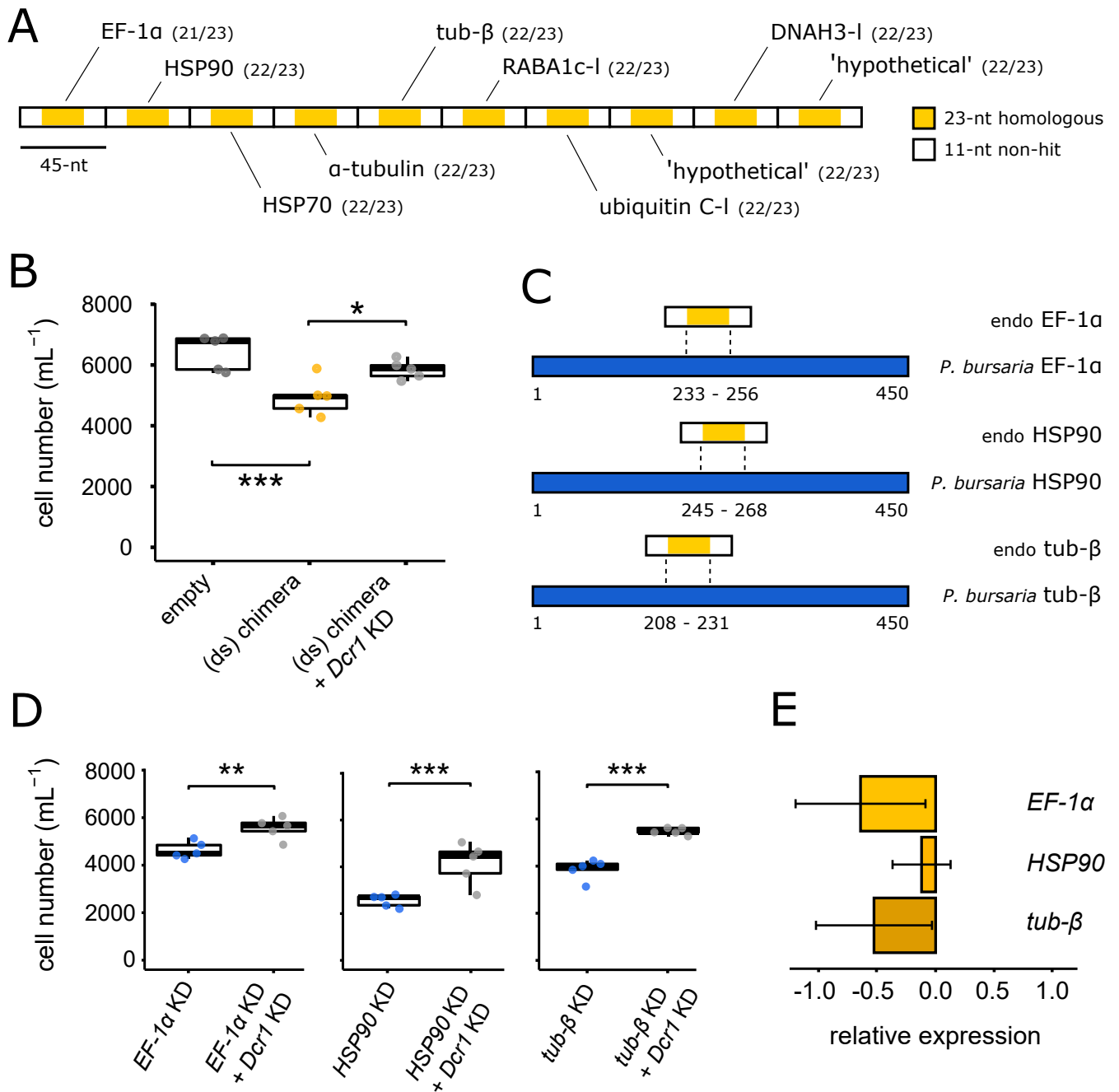
Fig. 3. In vivo exposure to a synthetic endosymbiont-derived RNA chimera generates simultaneous knockdown (KD) of homologous host genes in *P. bursaria*. (*A*) Schematic showing chimeric construct design. Ten endosymbiont-derived 45-nt transcript sequences (Dataset S3), featuring a 23-nt region with >91% sequence identity between endosymbiont and host (yellow) flanked by 11-nt of "nonhit" algal transcript (white). Numbers in brackets denote sequence identity between endosymbiont and host. (*B*) *P. bursaria* cell number after 12 d of feeding with *E. coli* expressing chimera dsRNA (yellow), chimera dsRNA mixed with Dicer (*Dcr1*) dsRNA (light gray; rescue), or an empty vector control (dark gray). See *SI Appendix*, Fig. S9 for an experimental dilution of dsRNA chimera delivery. (*C*) Schematic demonstrating the degree of overlap between three endosymbiont (endo)-derived transcripts from the chimeric construct (white/yellow) and respective 450-nt homologous region of the host (*P. bursaria*) transcript (blue). Each region of 23-nt overlap (yellow) represents putative RNAi "collisions" between endosymbiont and host. (*D*) *P. bursaria* cell number after 12 d of feeding with *E. coli* expressing host *EF1-α*, *HSP90*, or *tub-β* 450-nt dsRNA (blue) to induce host KD compared to Dicer (*Dcr1*) dsRNA mixed controls (gray; rescue phenotype). See also *SI Appendix*, Fig. S11 for individual KD of each of the 10 broader host targets of the endosymbiont-derived dsRNA chimera. (*E*) qPCR of mRNA extracted from day 3 of chimera-RNAi feeding (*B*), revealing KD of *EF-1α*, *HSP90*, and *tub-β* host gene expression in *P. bursaria* in response to endosymbiont-derived chimera dsRNA exposure. Standardized expression of an *Actin* housekeeping gene was used for normalization. Data are represented as mean ± SD of three biological replicates. These three genes were assessed for KD as a result of endosymbiont-derived chimera exposure, based on evidence that directed KD of the wider corresponding host gene led to retardation of *P. bursaria* culture growth (*D*). For an extended figure showing the effect of Dicer KD on the expression of these three host genes, see *SI Appendix*, Fig. S12. (*B* and *D*) Multiple vector delivery was conducted at a 50:50 ratio during feeding. Boxplot data are represented as maximum, upper quartile (Q3), mean, lower quartile (Q1) and minimum values of five biological replicates. Individual data points are shown. Significance for boxplot data were calculated as *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$ using a generalized linear model with quasi-Poisson distribution.

Single-Stranded Delivery of Synthetic Endosymbiont-Derived RNA, Analogous to Endosymbiont mRNA, Results in a Cost to *P. bursaria* Growth. As RNA derived naturally from the endosymbiont was unlikely to be double-stranded, two further constructs were designed to assess the efficacy of synthetic endosymbiont chimera ssRNA exposure in *P. bursaria* (Fig. 4 and *SI Appendix*, Fig. S13). Exposure to endosymbiont chimera ssRNA also resulted in a significant cost to *P. bursaria* culture growth; however, this effect was only observed when delivered in the sense orientation ([+]ssRNA). Notably, chimera [+]ssRNA exposure resulted in a greater cost to host growth than chimera dsRNA exposure. A similar orientation bias was observed upon Dicer knockdown during cycloheximide treatment in a prior experiment, in which delivery of Dicer [+]ssRNA rescued culture growth to a greater extent than delivery of Dicer [−]ssRNA or Dicer dsRNA (*SI Appendix*, Fig. S14). Importantly, the orientation of [+]ssRNA represents the same orientation as the mRNA transcripts from which each of these target templates were identified.

Recursive knockdown of host Dicer (dsRNA delivery) was able to attenuate the cost to host growth associated with endosymbiont chimera [+]ssRNA exposure (Fig. 4). Importantly, simultaneous knockdown of host *Rdr1* and *Rdr2* [RNA-dependent RNA polymerases involved in amplification of primary or secondary sRNA triggers during RNAi (43, 47)] also significantly rescued this cost to host growth. This is consistent with the role of RdRP proteins in processing exogenous sRNA, partially degraded mRNA cleavage products, or full-length mRNA transcripts in *Paramecium* (43, 45–47). As for Dicer, recursive knockdown of host RdRP through RNAi represents a potential paradox, and so constitutes only a partial loss of function (42, 52). However, we are again confident that disruption in this

manner would be sufficient to perturb RdRP-mediated RNAi in *P. bursaria* after 14 d, as seen in Fig. 4. While the mechanism of down-stream Dicer processing of RdRP-generated dsRNA substrates remains unknown (*SI Appendix*), we can nonetheless infer from these data that ssRNA-induced RNAi knockdown in *P. bursaria* is partially dependent upon host RdRP function.

Irrespective of the mechanistic basis for a sense-oriented ssRNA knockdown bias, it is important to note that delivery of chimera [+]ssRNA demonstrated here represents the same orientation as the endosymbiont-derived mRNA interaction fragments from which these putative endosymbiont–host RNA–RNA interactions were identified. This is consistent with the observation that *Paramecium* RNAi factors are capable of processing both endogenously and exogenously derived mRNA (43, 45). The ability to reduce host growth through delivery of ssRNA, in the same orientation as host mRNA, therefore supports the hypothesis that a cost to host growth upon endosymbiont digestion can be facilitated by host RNAi-mediated RNA–RNA interactions between endosymbiont and host mRNA populations.

Discussion

Through manipulation of the *P. bursaria* endosymbiotic system, we have demonstrated that RNA released upon digestion of the algal endosymbiont is processed by the host RNAi system. For endosymbiont-derived mRNA sharing a high level of sequence identity with host transcripts, this processing can result in knockdown of endogenous host gene expression, resulting in a cost to host growth (*SI Appendix*, Fig. S1). We therefore postulate that these RNA–RNA interactions are of importance in eukaryote–eukaryote endosymbioses where partners are likely to share greater sequence similarity, especially among conserved transcripts which tend to be more highly represented among lethal or conditionally essential genes (61). Due to the inherent difficulty in characterizing such mechanisms directly, we have relied on multiple-experimental approaches to demonstrate the viability of putative RNA–RNA interactions at each stage of the process. We have tracked the interaction through sRNA sequencing, recapitulated the effect through exposure to synthetic endosymbiont-derived RNA—including sense ssRNA analogous to endosymbiont mRNA—and demonstrated that this mechanism is mediated by host Dicer, AGO-Piwi, Pds1, and RdRP proteins. This process of host gene knockdown in response to endosymbiont-derived RNA processing by host RNAi factors, which we term "RNAi collisions," represents a mechanism to sanction the host for breakdown of the interaction, a factor that would promote stability in a facultative eukaryote–eukaryote endosymbiosis.

The long-term maintenance of symbiotic interactions represents a quandary for evolutionary theory (31–36). How do such relationships avoid overexploitation by one partner that would ultimately lead the interaction to collapse? Partner switching is one option; however, the result is typically a reduced pattern of coevolution between symbiont and host that can inhibit the process of metabolic and genetic integration, stalling the evolution of stable interactions which are needed if a system is to move toward evolution of an organelle (13, 62–65). Previous studies have suggested that when selfish behaviors arise, the evolution of enforcement to punish or suppress the exploitative partner can act to restore cooperation (9, 10). Enforcement mechanisms have been identified in diverse biological systems (66, 67), and are argued to be one of the most effective drivers of cooperation between individuals from different species. The results presented here suggest that "RNAi collisions" between endosymbiont and host, which are capable of imposing a cost to host growth for breakdown of the symbiosis, could provide an emergent mechanism to discourage overexploitation of the endosymbiont population by the host. Under such a scenario the host would gain greater benefit from sequestering the endosymbiont
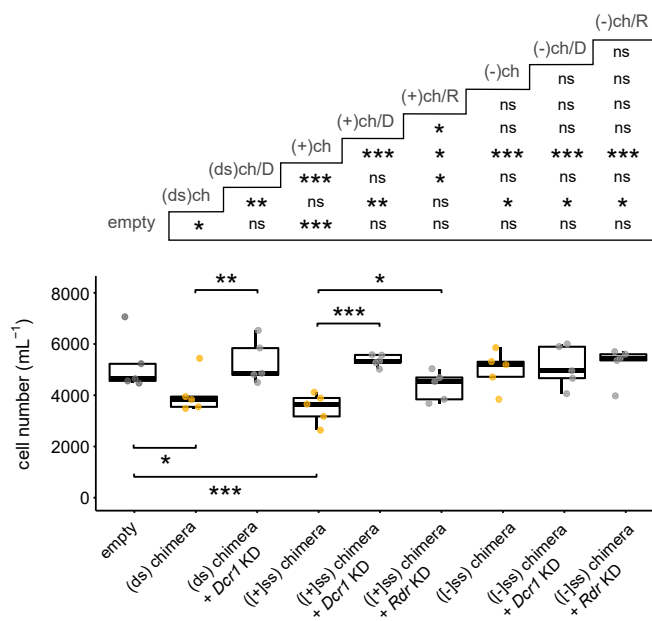

	empty	(ds)ch	(ds)ch/D	(+)ch	(+)ch/D	(+)ch/R	(-)ch	(-)ch/D	(-)ch/R
(-)ch/R									ns
(-)ch/D								ns	ns
(-)ch							ns	ns	ns
(+)ch/R					*	ns	ns	ns	ns
(+)ch/D				*	ns	ns	ns	ns	ns
(+)ch		***	***	***	*	***	***	***	
(ds)ch/D		**	***	ns	ns	*	ns	ns	ns
(ds)ch		*	**	ns	**	ns	*	*	*
empty		*	ns	***	ns	ns	ns	ns	ns

Fig. 4. Single-stranded delivery of synthetic endosymbiont-derived RNA, analogous to endosymbiont mRNA, results in a cost to *P. bursaria* growth. *P. bursaria* cell number after 14 d of feeding with *E. coli* expressing chimera RNA in ds, [+]ss or [−]ss orientation (yellow); chimera RNA in ds, [+]ss or [−]ss orientation mixed with Dicer (*Dcr1*) or RdRP (*Rdr*) dsRNA (light gray; rescue); or an empty vector control (dark gray). Multiple vector delivery was conducted at a 50:50 ratio during feeding. Asterisks displayed in the grid denote pairwise significance values. Boxplot data are represented as maximum, upper quartile (Q3), mean, lower quartile (Q1) and minimum values of five biological replicates. All significance for boxplot data was calculated as $*P \leq 0.05$, $**P \leq 0.01$, $***P \leq 0.001$, ns = no significance, using a generalized linear model with quasi-Poisson distribution.

Jenkins et al.
Emergent RNA–RNA interactions can promote stability in a facultative phototrophic endosymbiosis

PNAS | 7 of 11
https://doi.org/10.1073/pnas.2108874118

within a stable perialgal vacuole, protecting them from digestion and harvesting the nutrients which leak from them (26, 62), rather than simply digesting the would-be food through lysosomal fusion (14).

Interestingly, the emergence of this mechanism appears to be a by-product of preexisting biological features that are already likely to be under strong selective pressure. For instance, the essential RNAi system of the host, which in some ciliates can function in exogenous RNA processing, transposon elimination, nuclear rearrangement and transcriptional regulation (43–46, 48, 53, 68–71). Or, furthermore, the conserved gene repertoire and sequence composition of the host and endosymbiont transcriptomes from which some "RNAi collisions" have here been identified [including many transcripts which are conditionally essential in other systems (61)]. Unlike comparable mechanisms of RNA–RNA interactions that have been studied in host–pathogen symbioses (37–41), these "RNAi collisions" appear to be untargeted and, hence, emergent. In the aforementioned host–pathogen systems, targeted RNA is passed from one partner to the other in order to modulate expression of transcripts involved in virulence or resistance (37, 40, 41). However, in order for such systems to evolve, they must first exist in an untargeted form upon which selection is able to act, allowing the emergence of specific RNA factors (38, 39). Identification of undirected "RNAi collisions" in the *P. bursaria* system represents one such intermediary state, emergent in nature and untargeted, upon which sustained cellular interaction coupled with the potential for host-symbiont conflict could drive the selection of targeted RNA–RNA interactions.

We therefore propose that "RNAi collisions" represent a putative mechanism to discourage overexploitation of the endosymbiont population by the host. Here we have used the example of mass endosymbiont digestion in response to drug treatment to simulate this effect in the extreme. In natural interactions between *P. bursaria* and its algal endosymbiont, such a cost would only need to occur in the drastic occurrence of mass endosymbiont digestion in order to drive stability of the interaction. Importantly, the endosymbiotic algal population within *P. bursaria* is largely composed of closely related or clonal lineages (11–13), and as such, the fate of the algal population should be considered as a collective unit. This cost therefore need only act to suppress large-scale, rapid destruction by the host in order to drive the maintenance of a surviving subsection of the endosymbiont population. Previous studies have demonstrated how *P. bursaria* is capable of manipulating endosymbiont load in response to varying light conditions to better suit its own ends (28, 29), however, in these examples, reduction of endosymbiont number through digestion is slow and partial. By providing a system that selects against rapid and near-complete digestion of the endosymbiont population, "RNAi collisions" effectively buffer the endosymbiotic interaction against total breakdown. We suggest that this has allowed the relationship to be maintained across time and varying ecological conditions, even in the event of host-symbiont conflict (28, 29) and fluctuating endosymbiont numbers (14, 27, 28). As an alternate route to conflict resolution that avoids partner switching, we propose that such a mechanism would facilitate greater coevolution between endosymbiont and host. Over time, this would allow the metabolic and genetic integration that drives the formation of obligate symbioses to become manifest. We therefore present "RNAi collisions" as a mechanism in this endosymbiotic system, a factor that can promote stability in the face of conflict in an emergent endosymbiotic eukaryote–eukaryote cell–cell interaction.

Methods

Culture Conditions and Media. In all RNAi experiments, *P. bursaria* 186b (CCAP 1660/18) strain was used. For experiments requiring a nonphotoendosymbiotic *Paramecium* system for comparison, *P. tetraurelia* nd7 strain was used. For eDicer analysis, *P. bursaria* 186b and Yad1g1N strains were used (42, 60).

Paramecium cells were cultured in New Cereal Leaf Prescott Liquid media (NCL). NCL media was prepared by adding 4.3 mg · L^{-1} $CaCl_2.2H_2O$, 1.6 mg · L^{-1} KCl, 5.1 mg · L^{-1} K_2HPO_4, and 2.8 mg · L^{-1} $MgSO_4.7H_2O$ to deionized water. Then, 1 g · L^{-1} wheat bran was added, and the solution was boiled for 5 min. Once cooled, media was filtered once through Whatman Grade 1 filter paper and then through Whatman GF/C glass microfiber filter paper. Filtered NCL media was autoclaved at 121 °C for 30 min to sterilize prior to use.

NCL medium was bacterized with *Klebsiella pneumoniae* SMC and supplemented with 0.8 mg · L^{-1} β-sitosterol prior to propagation. *Paramecium* cells were subcultured 1:9 into fresh bacterized NCL media once per month for *P. bursaria* 186b and once every 2 wk for *P. tetraurelia* nd7. *Paramecium* cultures were maintained at 18 °C with a light–dark cycle of 12:12 h.

LysoTracker Staining and Fluorescent Imaging. *Paramecium* cells were fixed using 0.5% paraformaldehyde and incubated for 20 min at room temperature. Cells were then treated with 2 μM LysoTracker Green DND-26 (Invitrogen) to stain acidic vesicles such as the lysosome and incubated in constant darkness for 2 h at room temperature prior to imaging. Stained *Paramecium* cells were imaged on an ImageXpress Pico Automated Cell Imaging System at 10× magnification, using the Cy5 (absorbance, 630/40 nm; emission, 695/45 nm) and FITC (absorbance, 465/40 nm; emission, 525/30 nm) channels to capture algal-chlorophyll autofluorescence and LysoTracker Green DND-26 fluorescence, respectively. Captured images were automatically stitched together using the built-in CellReporterXpress acquisition and analysis software, generating a composite "tiled" image of the entire well. In some instances, minor distortion in fluorescence may be observed in cells close to the boundary of two "tiled" images as a result of this automated stitching process. Representative cell images were selected and cropped using the Snipping Tool to highlight the stages of endosymbiont elimination presented in Fig. 1.

Gene Synthesis and Construct Design. Sequences for plasmid constructs were synthesized de novo by either Genscript or SynBio Technologies and cloned into an L4440 plasmid vector. Sequences and cloning sites for each plasmid construct are detailed in Dataset S1. All modified constructs were confirmed by Sanger sequencing (Eurofins Genomics).

RNAi Feeding. *Paramecium* was fed with *E. coli* transformed with an L4440 plasmid construct with paired IPTG-inducible T7 promoters, facilitating targeted gene knockdown through the delivery of complementary dsRNA. L4440 plasmid constructs were transformed into *E. coli* HT115 competent cells and grown overnight on LB agar (50 μg · mL^{-1} ampicillin and 12.5 μg · mL^{-1} tetracycline) at 37 °C. Positive transformants were picked and grown overnight in LB (50 μg · mL^{-1} ampicillin and 12.5 μg · mL^{-1} tetracycline) at 37 °C with shaking (180 rpm). Overnight precultures were back-diluted 1:25 into 50 mL of LB (50 μg · mL^{-1} ampicillin and 12.5 μg · mL^{-1} tetracycline) and incubated for a further 2 h under the same conditions, until an OD_{600} of between 0.4 and 0.6 was reached. *E. coli* cultures were then supplemented with 0.4 mM IPTG to induce template expression within the L4440 plasmid and incubated for a further 3 h under the same conditions. *E. coli* cells were pelleted by centrifugation (3,100 × g for 2 min), washed with sterile NCL media, and pelleted once more. *E. coli* cells were then resuspended in NCL media supplemented with 0.4 mM IPTG, 100 μg · mL^{-1} ampicillin, and 0.8 μg · mL^{-1} β-sitosterol and adjusted to a final OD_{600} of 0.1.

Paramecium cells were pelleted by gentle centrifugation in a 96-well plate (10 min at 800 × g), taking care not to disturb the cell pellet by leaving 50 μL of supernatant, and resuspended 1:4 into 200 μL of induced *E. coli* culture media (to make 250 μL total). Feeding was conducted daily for up to 18 d using freshly prepared bacterized media.

Single-Stranded Construct Design and Confirmation. To create a version of L4440 which expressed only ssRNA, L4440 was digested with KpnI and PvuII (Promega), gel-purified (Wizard SV Gel and PCR Clean-Up System, Promega), and blunted using *PfuUltra* HF DNA polymerase (Agilent Technologies). The blunt vector was religated using T4 DNA ligase (Thermo Scientific) and confirmed by sequencing (Eurofins Genomics), generating plasmid pDM004 which contains only a single T7 promoter. Fragments were then excised from their respective L4440 plasmids or amplified by PCR (Q5 Polymerase; New England Biolabs) to swap the restriction sites, digested, and ligated into pDM004 to generate plasmids containing inserted fragments in sense [+] or antisense [−] orientation. Generated plasmids were transformed into HT115 *E. coli* for use in RNAi feeding experiments.

To confirm that these plasmid constructs generated only ssRNA, cultures of *E. coli* HT115 containing pDM005-1 ([−]ssRNA chimera) or pDM005-2 ([+]ssRNA chimera) were grown overnight at 37 °C and 180 rpm in LB supplemented with 50 µg · mL^{-1} ampicillin and 12.5 µg · mL^{-1} tetracycline. Cultures were diluted 1:25 in fresh medium and grown to an OD$_{600}$ of 0.4 to 0.6. Expression was then induced with 400 µM IPTG for 3 h, after which 1 mL of culture was pelleted by centrifugation for 2 min at 3,100 × g. RNA was extracted using an RNeasy Mini Kit (Qiagen), following the manufacturer's protocol for Total RNA Purification from Animal Cells. The 20 µL of RNA was then treated with 10 µg · mL^{-1} RNase A (Sigma-Aldrich) for 1 h at 30 °C in the presence of 300 mM NaCl [stabilizing dsRNA and allowing RNase A to degrade only ssRNA (72)]. A separate aliquot was left untreated, with 300 mM NaCl added to facilitate precipitation. All samples were then extracted with 1:1 phenol:chloroform and precipitated with two volumes of ethanol. Pellets were washed twice with 80% ethanol and resuspended in nuclease-free water. RNA samples were then cleared of residual genomic DNA using the TURBO DNA-*free* Kit (Ambion), following the manufacturer's protocol for routine DNase treatment. RT-PCR was then performed using the Qiagen OneStep RT-PCR kit following the manufacturer's instructions, with 0.5 µL template RNA and 0.6 µM each primer (pDM005_RT_F: 5′-ACTTCAATGATT CGCAGCGG-3′ and pDM005_RT_R: 5′-AAGTAGCTGCTGTTCTCGGT-3′), generating an 85-nt PCR product. Cycling conditions were as detailed in the manufacturer's protocol (30 cycles), with 1 min annealing at 50 °C. PCR products were then resolved on a 2% agarose gel to assess for the presence/absence of amplification in each sample (*SI Appendix*, Fig. S13).

qPCR Analysis. RNA was extracted from *P. bursaria* 186b for gene expression analysis after 3 d of RNAi feeding. *Paramecium* cells (∼10^3 per culture) were pelleted by gentle centrifugation (800 × g for 10 min), snap-frozen in liquid nitrogen, and stored at −80 °C. RNA extraction was performed using TRIzol reagent (Invitrogen), following the manufacturer's protocol after resuspending each pellet in 900 µL TRIzol reagent. RNA was precipitated using GlycoBlue Coprecipitant (Invitrogen) to aid RNA pellet visualization and then cleared of residual DNA using the TURBO DNA-*free* Kit (Ambion), following the manufacturer's protocol for routine DNase treatment.

RNA was reverse transcribed into single-stranded cDNA using the SuperScript III First-Strand Synthesis SuperMix (Invitrogen), following the manufacturer's protocol. qPCR was performed in a StepOnePlus Real-Time PCR system (Thermo Fisher Scientific). Reaction conditions were optimized using a gradient PCR, with a standard curve determined using 10-fold dilutions of *P. bursaria* cDNA: *EF-1α* (slope: −3.353; R^2: 0.999; efficiency: 98.740%), *HSP90* (slope: −3.319; R^2: 0.998; efficiency: 100.131%), *tub-β* (slope: −3.378; R^2: 0.992; efficiency: 97.692%), and *actin* (slope: −3.349; R^2: 0.983; efficiency: 98.866%), using StepOne software v2.3. Each 20 µL reaction contained 10 µL PowerUp SYBR Green Master Mix (Thermo Fisher Scientific), 500 nM each primer (300 nM for *tub-β*), and 1 µL (50 ng) cDNA. Each reaction was performed in duplicate for each of three biological replicates, alongside a "no-RT" (i.e., nonreverse transcribed RNA) control to detect any genomic DNA contamination. Cycling conditions were as follows: UDG activation, 2 min at 50 °C and DNA polymerase activation, 2 min at 95 °C, followed by 40 cycles of 15 s, 95 °C and 1 min at 55 to 65 °C (*EF-1α* [58 °C], *HSP90* [58 °C], *tub-β* [57 °C], and *actin* [58 °C]). Primer pairs for each reaction are listed in Dataset S5. Each reaction was followed by melt-curve analysis, with a 60 to 95 °C temperature gradient (0.3 °C · s^{-1}), ensuring the presence of only a single amplicon, and ROX was used as a reference dye for calculation of C$_T$ values. C$_T$ values were then used to calculate the change in gene expression of the target gene in RNAi samples relative to control samples, using a derivation of the 2$^{-\Delta\Delta CT}$ algorithm (73).

sRNA Isolation and Sequencing. Total RNA for sRNA sequencing was extracted from *P. bursaria* (or free-living algal) cultures using TRIzol reagent (Invitrogen), as detailed previously. To isolate sRNA from total RNA, samples were size separated on a denaturing 15% TBE-UREA polyacrylamide gel. Gels were prepared with a 15-mL mix with final concentrations of 15% Acrylamide/Bis (19:1), 8M UREA, TBE (89 mM Tris, 89 mM Borate, 2 mM EDTA), and the polymerization started by the addition of 150 µL 10% APS (Sigma-Aldrich) and 20 µL TEMED (Sigma-Aldrich). Gels were pre-equilibrated by running for 15 min (200 V, 30 mA) in TBE before RNA loading. The ladder mix consisted of 500 ng ssRNA ladder (50- to 1,000-nt, NEB#N0364S) and 5 to 10 ng of each 21- and 26-nt RNA oligo loaded per lane. The marker and samples were mixed with 2× RNA loading dye (NEB) and heat denatured at 90 °C for 3 min before snap cooling on ice for 2 min prior to loading. Blank lanes were left between samples/replicates to prevent cross-contamination during band excision. Gels were then run for 50 min (200V, 30 mA).

Once run, gels were stained by shaking (60 rpm) for 20 min at room temperature in a 40-mL TBE solution containing 4 µL SYBR Gold Nucleic Acid Gel Stain. Bands of the desired size range (∼15- to 30-nt) were visualized under blue light, excised, and placed into a 0.5-mL tube pierced at the bottom by a 21-gauge needle, resting within a 1.5-mL tube, and centrifuged (16,000 × g for 1 min). Then, 400 µL of RNA elution buffer (1 M sodium acetate pH 5.5 and 1 mM EDTA) was added to the 1.5-mL tube containing centrifuged gel slurry, and the empty 0.5-mL tube was discarded. Gel slurry was manually homogenized until dissolved using a 1-mL sterile plunger and incubated at room temperature for 2 h with shaking at 1,400 rpm.

Solutions containing RNA elution buffer and gel slurry were transferred to a Costar Spin-X 0.22 µm filter column and centrifuged (16,000 × g for 1 min). The filter insert containing acrylamide was discarded. Then, 1 mL of 100% EtOH was added to each solution, alongside 15 µg of GlycoBlue Coprecipitant (Invitrogen), to aid sRNA pellet visualization, and stored overnight at −80 °C to precipitate. Precipitated solutions were centrifuged at 4 °C (12,000 × g for 30 min), and the supernatant was discarded. sRNA pellets were washed with 500 µL of cold 70% EtOH (12,000 × g for 15 min at 4 °C) and air dried in a sterile PCR hood for 10 min before resuspending in 15 µL of RNase-free water and storage at −80 °C.

sRNA-Sequencing and Read Processing. sRNA concentrations were determined using an Agilent 2,100 Bioanalyzer, following the Agilent Small RNA kit protocol, and all samples matched to 0.7 ng · mL^{-1} prior to sequencing. Library preparation and subsequent RNA-sequencing (RNA-seq) was performed for 54 samples using 50-bp paired-end rapid run across four lanes on an Illumina HiSEq 2500, yielding ∼120 to 150 million paired-end reads per lane (∼9 to 11 million paired-end reads per sample).

The raw paired-end reads from the RNA-seq libraries were trimmed using Trim Galore in order to remove barcodes (4-nt from each 3′- and 5′-end) and sRNA adaptors, with additional settings of a phred-score quality threshold of 20 and minimum length of 16-nt. Result were subsequently checked with FastQC.

Assigning sRNAs to the Algal Endosymbiont Transcript Bins. Trimmed reads were mapped against the "endosymbiont" dataset of assembled transcripts using the HISAT2 alignment program with default settings. Postmapping, the BAM files were processed using SAMTOOLS and a set of custom scripts (https://github.com/guyleonard/paramecium) to produce a table of mapped read accessions and their respective read lengths. Using these exported count tables, of mapped reads per read length, transcripts with >10 hits for read lengths between 21- to 25-nt were searched using reciprocal BLASTX against the National Center for Biotechnology Information (NCBI) nonredundant "nr" proteins sequence database in order to assign taxonomic identity to each transcript. This allowed filtering of the main "endosymbiont" dataset into subsets corresponding to either algal cytoplasmic mRNA, algal cytoplasmic rRNA, algal plastid RNA, algal mitochondrial RNA, host RNA contamination, bacterial RNA contamination, or vector RNA contamination. All identified algal cytoplasmic rRNA (28 transcripts), plastid RNA (56 transcripts), or mitochondrial RNA (18 transcripts) sequences were sorted into new bins representing each RNA "species." Host (5 transcripts) and bacterial (34 transcripts) contamination were sorted into respective "host" or "bacterial" datasets. Vector (1 transcript) and all unidentifiable sequences (20 transcripts) were separated into a dataset labeled as "other." All remaining transcripts with >10 sRNA hits for read sizes 21- to 25-nt in the original dataset now represented putative algal mRNA transcripts (i.e., for 23-nt reads, this represented 3,659 sRNA reads mapping to 148 transcripts). In this algal mRNA dataset, we also included transcripts that fell below the >10 sRNA hit threshold for manual curation, and which therefore may correspond to reads mapping to either algal, host, bacterial, or "other" bins (i.e., for 23-nt reads, this represented 1,949 reads mapping to 605 transcripts in total, of which 1,408 reads were mapping to 468 transcripts which were not manually curated and so could be affected by contamination). This manual binning process was carried out in order to double check the original automated binning posttranscriptome assembly and to check for possible chimeric host-algal transcript sequences produced as a by-product of cDNA synthesis and transcriptome sequencing and assembly. This process collectively allowed accurate segregation of the existing "endosymbiont" transcript bin into discrete subsets based on RNA "species."

Using the HISAT2 alignment program with default settings, trimmed reads were mapped once more against the newly filtered algal mRNA, rRNA, plastid, and mitochondrial datasets. Count tables of mapped reads per read length were once again generated from the BAM files and used to plot a size distribution of 21- to 29-nt endosymbiont-derived sRNA abundance per RNA species. Size distributions of sRNA abundance for each sample were plotted

Jenkins et al.
Emergent RNA–RNA interactions can promote stability in a facultative phototrophic endosymbiosis

EVOLUTION

using the R programming language packages tidyverse, grid.extra, and ggplot2 in R Studio.

eDicer Methods for Identifying Putative RNA–RNA Interactions. To predict putative mRNA–mRNA interactions using eDicer (54), both the "host" and "endosymbiont" transcript bins processed from transcriptome data for *P. bursaria* Yad1g1N (60) were further filtered to minimize the risk of false positives for host and endosymbiont cross-comparisons. For the "host" bins, any transcript with >90% ID BLASTN hit to the following genome assemblies were removed: *Chlamydomonas reinhardtii* cc503 cw92 mt, *C. reinhardtii* v3, *Chlorella sorokiniana* 1228 v2, *C. sorokiniana* DOE1412 v3, or *C. sorokiniana* utex1230 lanl v2 assemblies from Los Alamos National Labs Greenhouse genome database. Similarly, for the "endosymbiont" bins, any transcript with >90% ID BLASTN hit to the following ParameciumDB genome assemblies were removed: *P. biaurelia* V1-4 v1, *P. caudatum* 43c3d v1, *P. bursaria* MAC 110224 v1, *P. decaurelia* MAC 223 v1, *P. dodecaurelia* MAC 274 v1, *P. jenningsi* MAC M v1, *P. novaurelia* MAC TE v1, *P. octaurelia* K8 CA1, *P. quadecaurelia* MAC NiA v1, *P. primaurelia* Ir4-2 v1, *P. tetraurelia* MAC 51 with and without IES, or *P. tetraurelia* MAC. Additionally, the reference (refseq) genome of *E. coli* MG1655 (NCBI: NZ_CP012868.1) and *K. pneumoniae* HS11286 (NCBI: NC_016845.1) were used as "food" comparison datasets.

The filtered "endosymbiont" transcript bin, *E. coli* CDS and *K. pneumoniae* CDS were decomposed into all possible 21- to 23-nt reads with jellyfish v2.2.10 and aligned with 95% identity to the "host" bin using Bowtie v1.2.3 (recommended for short alignments). These in silico RNA–RNA interaction simulations were performed using a wrapper tool we created named eDicer (v1.0.0) (54). The number of distinct aligning k-mers (i.e., putative interactions of unique sequences) for each pair of bins was then normalized by dividing by the total number of distinct k-mers in both bins, and ×100, to calculate a Jaccard Index % (i.e., normalized set similarity).

To predict putative "lethal" RNA–RNA interactions, a separate analysis was conducted using a subset of each dataset that was putatively homologous to a yeast "lethal" gene database (61). This "lethal" dataset contained genes known to be conditionally essential in *S. cerevisiae*. These putative "lethal" homologs for each dataset were identified using a tBLASTx search of the yeast "lethal" database (61) with a gathering threshold set at 1e-10 with a minimum of 50% sequence identity. Once curated, each "lethal" dataset was subject to eDicer analysis as described.

To predict putative rRNA–rRNA interactions using eDicer, a further analysis was conducted using a dataset consisting of full-length rRNA clusters for *Microactinium conductrix* [NCBI: ASM224581v2 (74)], *K. pneumoniae* [NCBI: NC_016845.1 (75)], *E. coli* (NCBI: NZ_CP012868.1), *P. bursaria* Yad1g1N (60), and *P. bursaria* 186b. Once again, each dataset was subject to eDicer analysis as described. All *in-silico* RNA–RNA interaction predictions were plotted using the R programming language packages tidyverse, grid.extra, and ggplot2 in R Studio.

Manual Curation of Additional Host Transcript Bins. In order to identify host transcripts that could be impacted by putative RNA–RNA interactions as a result of endosymbiont-derived sRNA exposure, trimmed Illumina reads for all sRNA sequencing samples were mapped to the "endosymbiont mRNA" transcript dataset. Resulting mapping files were filtered with custom scripts (76) producing tables of mapped hits with their respective read lengths. Reads from all tables of length 23-nt only were then extracted with SEQTK (resulting in 3,690 total). A BLASTn search of these 23-nt endosymbiont reads was conducted against the "host" transcript dataset to identify host transcripts with ≥95% identity over a 23-nt region. This resulted in three candidate host transcripts. These three host transcripts were searched using BLASTx against the NCBI nonredundant "nr" protein database, resulting in one transcript which had a 23-bp region with 100% sequence identity to both endosymbiont and host *HSP90*. For the remaining two transcripts, SNPs present in the identified 23-nt mapped sequence indicate that these are putative host sequences with 2-nt of mismatch compared to the respective endosymbiont sequence and were therefore unlikely to represent a product of RNA–RNA interactions between endosymbiont and host as a result of endosymbiont-derived 23-nt sRNA exposure.

In order to identify host transcripts that could be classified as "nonhit" transcripts (i.e., host transcripts sharing a low level of sequence identity with algal transcripts over 23-nt regions), the "host" dataset was searched using an organism-specific BLASTn (Chlorellaceae [NCBI: taxid35461]) against the NCBI nucleotide "nr/nt" database with a minimum expectation of 1e-05. Any host transcripts sharing >20-nt sequence identity with algal transcripts over a 23-nt region were rejected. This process was repeated until a dataset of 20 "non-hit-algal" host transcripts were identified. These putative "non-hit-algal" host transcripts were then searched using BLASTx against the NCBI nonredundant "nr" protein database to confirm that these 20 transcripts were derived from *Paramecium* and were not bacterial contamination.

1. J. M. Archibald, Endosymbiosis and eukaryotic cell evolution. *Curr. Biol.* **25**, R911–R921 (2015).
2. D. Baurain et al., Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. *Mol. Biol. Evol.* **27**, 1698–1709 (2010).
3. P. J. Keeling, The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annu. Rev. Plant Biol.* **64**, 583–607 (2013).
4. J. M. Archibald, The puzzle of plastid evolution. *Curr. Biol.* **19**, R81–R88 (2009).
5. C. J. Howe, A. C. Barbrook, R. E. R. Nisbet, P. J. Lockhart, A. W. D. Larkum, The origin of plastids. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 2675–2685 (2008).
6. J. N. Timmis, M. A. Ayliffe, C. Y. Huang, W. Martin, Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* **5**, 123–135 (2004).
7. M. W. Gray, Mitochondrial evolution. *Cold Spring Harb. Perspect. Biol.* **4**, a011403 (2012).
8. N. A. Moran, Symbiosis as an adaptive process and source of phenotypic complexity. *Proc. Natl. Acad. Sci. U.S.A.* **104** (suppl. 1), 8627–8633 (2007).
9. J. A. Ågren, N. G. Davies, K. R. Foster, Enforcement is central to the evolution of cooperation. *Nat. Ecol. Evol.* **3**, 1018–1029 (2019).
10. S. A. West, E. T. Kiers, E. L. Simms, R. F. Denison, Sanctions and mutualism stability: Why do rhizobia fix nitrogen? *Proc. Biol. Sci.* **269**, 685–694 (2002).
11. U. E. Achilles-Day, J. G. Day, Isolation of clonal cultures of endosymbiotic green algae from their ciliate hosts. *J. Microbiol. Methods* **92**, 355–357 (2013).
12. R. Hoshina, Y. Kusuoka, DNA analysis of algal endosymbionts of ciliates reveals the state of algal integration and the surprising specificity of the symbiosis. *Protist* **167**, 174–184 (2016).
13. P. Zagata, M. Greczek-Stachura, S. Tarcz, M. Rautian, The evolutionary relationships between endosymbiotic green algae of *Paramecium bursaria* syngens originating from different geographical locations. *Folia Biol. (Krakow)* **64**, 47–54 (2016).
14. Y. Kodama, M. Fujishima, Cycloheximide induces synchronous swelling of perialgal vacuoles enclosing symbiotic *Chlorella vulgaris* and digestion of the algae in the ciliate *Paramecium bursaria*. *Protist* **159**, 483–494 (2008).
15. Y. Kodama, M. Fujishima, Symbiotic *Chlorella variabilis* incubated under constant dark conditions for 24 hours loses the ability to avoid digestion by host lysosomal enzymes in digestive vacuoles of host ciliate *Paramecium bursaria*. *FEMS Microbiol. Ecol.* **90**, 946–955 (2014).
16. G. Omura et al., A bacteria-free monoxenic culture of *Paramecium bursaria*: Its growth characteristics and the re-establishment of symbiosis with *Chlorella* in bacteria-free conditions. *Jpn. J. Protozool* **37**, 139–150 (2004).
17. M. Tanaka et al., Complete elimination of endosymbiotic algae from *Paramecium bursaria* and its confirmation by diagnostic PCR. *Acta Protozool.* **41**, 255–261 (2002).
18. J. A. Brown, P. J. Nielsen, Transfer of photosynthetically produced carbohydrate from endosymbiotic Chlorellae to *Paramecium bursaria*. *J. Protozool.* **21**, 569–570 (1974).
19. G. F. Esteban, T. Fenchel, B. J. Finlay, Mixotrophy in ciliates. *Protist* **161**, 621–641 (2010).
20. M. D. Johnson, Acquired phototrophy in ciliates: A review of cellular interactions and structural adaptations. *J. Eukaryot. Microbiol.* **58**, 185–195 (2011).
21. Y. Kato, N. Imamura, Amino acid transport systems of Japanese *Paramecium* symbiont F36-ZK. *Symbiosis* **47**, 99–107 (2009).
22. Y. Kato, N. Imamura, Effect of sugars on amino acid transport by symbiotic *Chlorella*. *Plant Physiol. Biochem.* **46**, 911–917 (2008).
23. H. Kawakami, N. Kawakami, Behavior of a virus in a symbiotic system, *Paramecium bursaria*—Zoochlorella. *J. Protozool.* **25**, 217–225 (1978).
24. R. C. Parker, Symbiosis in *Paramecium bursaria*. *J. Exp. Zool.* **46**, 1–12 (1926).

25. E. Ziesenisz, W. Reisser, W. Wiessner, Evidence of de novo synthesis of maltose excreted by the endosymbiotic *Chlorella* from *Paramecium bursaria*. *Planta* **153**, 481–485 (1981).

26. M. E. S. Sørensen *et al*., Comparison of independent evolutionary origins reveals both convergence and divergence in the metabolic mechanisms of symbiosis. *Curr. Biol.* **30**, 328–334.e4 (2020).

27. Y. Kodama, M. Fujishima, Infectivity of *Chlorella* species for the ciliate *Paramecium bursaria* is not based on sugar residues of their cell wall components, but on their ability to localize beneath the host cell membrane after escaping from the host digestive vacuole in the early infection process. *Protoplasma* **231**, 55–63 (2007).

28. C. D. Lowe, E. J. Minter, D. D. Cameron, M. A. Brockhurst, Shining a light on exploitative host control in a photosynthetic endosymbiosis. *Curr. Biol.* **26**, 207–211 (2016).

29. M. E. S. Sørensen *et al*., The role of exploitation in the establishment of mutualistic microbial symbioses. *FEMS Microbiol. Lett.* **366**, fnz148 (2019).

30. M. E. S. Sørensen, A. J. Wood, D. D. Cameron, M. A. Brockhurst, Rapid compensatory evolution can rescue low fitness symbioses following partner-switching. *bioRxiv* [Preprint] (2020). https://doi.org/10.1101/2020.11.06.371401 (Accessed 8 March 2021).

31. R. Axelrod, W. D. Hamilton, The evolution of cooperation. *Science* **211**, 1390–1396 (1981).

32. K. R. Foster, T. Wenseleers, A general model for the evolution of mutualisms. *J. Evol. Biol.* **19**, 1283–1293 (2006).

33. S. A. Frank, *Foundations of Social Evolution* (Princeton University Press, Princeton, NJ, 1998).

34. E. A. Herre, N. Knowlton, U. G. Mueller, S. A. Rehner, The evolution of mutualisms: Exploring the paths between conflict and cooperation. *Trends Ecol. Evol.* **14**, 49–53 (1999).

35. E. Szathmáry, J. M. Smith, The major evolutionary transitions. *Nature* **374**, 227–232 (1995).

36. S. A. West, R. M. Fisher, A. Gardner, E. T. Kiers, Major evolutionary transitions in individuality. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10112–10119 (2015).

37. A. H. Buck *et al*., Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity. *Nat. Commun.* **5**, 5488 (2014).

38. Z. Guo, Y. Li, S. W. Ding, Small RNA-based antimicrobial immunity. *Nat. Rev. Immunol.* **19**, 31–44 (2019).

39. C. Y. Huang, H. Wang, P. Hu, R. Hamby, H. Jin, Small RNAs - big players in plant-microbe interactions. *Cell Host Microbe* **26**, 173–182 (2019).

40. M. Wang *et al*., Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. *Nat. Plants* **2**, 16151 (2016).

41. A. Weiberg *et al*., Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* **342**, 118–123 (2013).

42. B. H. Jenkins *et al*., Characterization of the RNA-interference pathway as a tool for reverse genetic analysis in the nascent phototrophic endosymbiosis, *Paramecium bursaria*. *R. Soc. Open Sci.* **8**, 210140 (2021).

43. Q. Carradec *et al*., Primary and secondary siRNA synthesis triggered by RNAs from food bacteria in the ciliate *Paramecium tetraurelia*. *Nucleic Acids Res.* **43**, 1818–1833 (2015).

44. A. Galvani, L. Sperling, RNA interference by feeding in *Paramecium*. *Trends Genet.* **18**, 11–12 (2002).

45. S. Karunanithi *et al*., Exogenous RNAi mechanisms contribute to transcriptome adaptation by phased siRNA clusters in *Paramecium*. *Nucleic Acids Res.* **47**, 8036–8049 (2019).

46. S. Marker, Q. Carradec, V. Tanty, O. Arnaiz, E. Meyer, A forward genetic screen reveals essential and non-essential RNAi factors in *Paramecium tetraurelia*. *Nucleic Acids Res.* **42**, 7268–7280 (2014).

47. S. Marker, A. Le Mouël, E. Meyer, M. Simon, Distinct RNA-dependent RNA polymerases are required for RNAi triggered by double-stranded RNA versus truncated transgenes in *Paramecium tetraurelia*. *Nucleic Acids Res.* **38**, 4092–4107 (2010).

48. F. Drews *et al*., Two Piwis with Ago-like functions silence somatic genes at the chromatin level. *bioRxiv* [Preprint] (2020). https://doi.org/10.1101/2020.08.24.263970 (Accessed 3 January 2021).

49. S. Karunanithi *et al*, Feeding exogenous dsRNA interferes with endogenous sRNA accumulation in Paramecium. *DNA Res.* **27**, dsaa005 (2020).

50. H. Cerutti, X. Ma, J. Msanne, T. Repas, RNA-mediated silencing in algae: Biological roles and tools for analysis of gene function. *Eukaryot. Cell* **10**, 1164–1172 (2011).

51. Y. Qiao *et al*., Small RNAs in plant immunity and virulence of filamentous pathogens. *Annu. Rev. Phytopathol.*, 10.1146/annurev-phyto-121520-023514 (2021).

52. W. F. Marshall, Modeling recursive RNA interference. *PLoS Comput. Biol.* **4**, e1000183 (2008).

53. C. Hoehener, I. Hug, M. Nowacki, Dicer-like enzymes with sequence cleavage preferences. *Cell* **173**, 234–247.e7 (2018).

54. F. Maguire, eDicer. GitHub. https://github.com/fmaguire/eDicer. Deposited 14 August 2020.

55. F. Ahmed, G. P. S. Raghava, Designing of highly effective complementary and mismatch siRNAs for silencing a gene. *PLoS One* **6**, e23443 (2011).

56. Q. Du, H. Thonberg, J. Wang, C. Wahlestedt, Z. Liang, A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. *Nucleic Acids Res.* **33**, 1671–1677 (2005).

57. A. L. Jackson *et al*., Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA* **12**, 1179–1187 (2006).

58. J. G. Doench, C. P. Petersen, P. A. Sharp, siRNAs can function as miRNAs. *Genes Dev.* **17**, 438–442 (2003).

59. S. L. Ameres, J. Martinez, R. Schroeder, Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* **130**, 101–112 (2007).

60. Y. Kodama *et al*., Comparison of gene expression of *Paramecium bursaria* with and without *Chlorella variabilis* symbionts. *BMC Genomics* **15**, 183 (2014).

61. J. A. Cotton, J. O. McInerney, Eukaryotic genes of archaebacterial origin are more important than the more numerous eubacterial genes, irrespective of function. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 17252–17255 (2010).

62. M. Fujishima, Y. Kodama, Endosymbionts in *Paramecium*. *Eur. J. Protistol.* **48**, 124–137 (2012).

63. J. L. Matthews *et al*., Partner switching and metabolic flux in a model cnidarian-dinoflagellate symbiosis. *Proc. Biol. Sci.* **285**, 20182336 (2018).

64. Y. Matsuura *et al*., Recurrent symbiont recruitment from fungal parasites in cicadas. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E5970–E5979 (2018).

65. F. Husnik, J. P. McCutcheon, Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5416–E5424 (2016).

66. D. A. Hickey, Selfish DNA: A sexually-transmitted nuclear parasite. *Genetics* **101**, 519–531 (1982).

67. E. T. Kiers, R. A. Rousseau, S. A. West, R. F. Denison, Host sanctions and the legume-rhizobium mutualism. *Nature* **425**, 78–81 (2003).

68. G. Lepère *et al*., Silencing-associated and meiosis-specific small RNA pathways in *Paramecium tetraurelia*. *Nucleic Acids Res.* **37**, 903–915 (2009).

69. P. Y. Sandoval, E. C. Swart, M. Arambasic, M. Nowacki, Functional diversification of Dicer-like proteins and small RNAs required for genome sculpting. *Dev. Cell* **28**, 174–188 (2014).

70. C. D. Malone, A. M. Anderson, J. A. Motl, C. H. Rexer, D. L. Chalker, Germ line transcripts are processed by a Dicer-like protein that is essential for developmentally programmed genome rearrangements of *Tetrahymena thermophila*. *Mol. Cell. Biol.* **25**, 9151–9164 (2005).

71. I. V. Nekrasova, Н. И. Владимировна, А. А. Potekhin, П. А. Анатольевич, Diversity of RNA interference pathways in regulation of endogenous and exogenous sequences expression in ciliates *Tetrahymena* and *Paramecium*. *Ekol. Genet.* **17**, 113–125 (2019).

72. D. R. Hansen, N. K. Van Alfen, K. Gillies, W. A. Powell, Naked dsRNA associated with hypovirulence of *Endothia parasitica* is packaged in fungal vesicles. *J. Gen. Virol.* **66**, 2605–2614 (1985).

73. J. D. Zhang, R. Biczok, M. Ruschhaupt, ddCt: The ddCt algorithm for the analysis of quantitative real-time PCR (qRT-PCR). Bioconductor version 3.11. https://doi.org/10.18129/B9.bioc.ddCt. Accessed 9 September 2021.

74. M. B. Arriola *et al*., Genome sequences of *Chlorella sorokiniana* UTEX 1602 and *Micractinium conductrix* SAG 241.80: Implications to maltose excretion by a green alga. *Plant J.* **93**, 566–586 (2018).

75. P. Liu *et al*., Complete genome sequence of *Klebsiella pneumoniae* subsp. pneumoniae HS11286, a multidrug-resistant strain isolated from human sputum. *J. Bacteriol.* **194**, 1841–1842 (2012).

76. G. Leonard, Paramecium. GitHub. https://github.com/guyleonard/paramecium. Deposited 26 March 2020.

77. B. H. Jenkins *et al*., Supplemental data for "Emergent RNA-RNA interactions can promote cooperation in a facultative phototrophic endosymbiosis." *Figshare*. https://doi.org/10.6084/m9.figshare.c.4978160.v3. Deposited 24 August 2021.

EVOLUTION

Jenkins et al.
Emergent RNA–RNA interactions can promote stability in a facultative phototrophic endosymbiosis

PNAS | 11 of 11
https://doi.org/10.1073/pnas.2108874118