PNAS RESEARCH ARTICLE | ENVIRONMENTAL SCIENCES

Temperature dependence of parasitoid infection and abundance of a diatom revealed by automated imaging and classification

Dylan Catlett[a,1], Emily E. Peacock[a], E. Taylor Crockford[a], Joe Futrelle[a], Sidney Batchelder[a], Bethany L. F. Stevens[a], Rebecca J. Gast[a], Weifeng G. Zhang[b], and Heidi M. Sosik[a]

Edited by Edward DeLong, University of Hawai'i at Manoa, Honolulu, HI; received February 27, 2023; accepted May 19, 2023

Diatoms are a group of phytoplankton that contribute disproportionately to global primary production. Traditional paradigms that suggest diatoms are consumed primarily by larger zooplankton are challenged by sporadic parasitic "epidemics" within diatom populations. However, our understanding of diatom parasitism is limited by difficulties in quantifying these interactions. Here, we observe the dynamics of *Cryothecomonas aestivalis* (a protist) infection of an important diatom on the Northeast U.S. Shelf (NES), *Guinardia delicatula*, with a combination of automated imaging-in-flow cytometry and a convolutional neural network image classifier. Application of the classifier to >1 billion images from a nearshore time series and >20 survey cruises across the broader NES reveals the spatiotemporal gradients and temperature dependence of *G. delicatula* abundance and infection dynamics. Suppression of parasitoid infection at temperatures <4 °C drives annual cycles in both *G. delicatula* infection and abundance, with an annual maximum in infection observed in the fall-winter preceding an annual maximum in host abundance in the winter-spring. This annual cycle likely varies spatially across the NES in response to variable annual cycles in water temperature. We show that infection remains suppressed for ~2 mo following cold periods, possibly due to temperature-induced local extinctions of the *C. aestivalis* strain(s) that infect *G. delicatula*. These findings have implications for predicting impacts of a warming NES surface ocean on *G. delicatula* abundance and infection dynamics and demonstrate the potential of automated plankton imaging and classification to quantify phytoplankton parasitism in nature across unprecedented spatiotemporal scales.

diatom | protistan parasite | Imaging FlowCytobot | machine learning | *Guinardia delicatula*

Significance

Diatoms are unicellular algae whose "blooms" are associated with high primary productivity, prolific fisheries, and carbon flux to the deep ocean. Despite its potential impact on marine food webs, diatom parasitism is poorly understood due to challenges observing its prevalence and environmental controls at appropriate spatiotemporal scales. Here, we use automated plankton imaging and machine learning classification to elucidate abundance and parasitic infection dynamics of a biomass-dominant diatom on the Northeast U.S. Shelf (NES). We suggest that temperature indirectly regulates diatom abundance via direct suppression of parasitism. This temperature dependence implies that ongoing warming may enable parasitic infection to occur throughout the year, driving dramatic shifts in this diatom's abundance dynamics with potential cascading effects on the NES ecosystem.

Diatoms are a major source of oceanic primary production and play an outsized role in marine ecosystems and global biogeochemistry (1, 2). Diatoms frequently accumulate in response to nutrient enrichment of the surface ocean, and these "blooms" likely result in productive marine food webs and fisheries (1, 3). Traditional oceanographic paradigms explain this association by arguing that diatom production is predominantly consumed by large micro- and meso-zooplankton which provide an efficient link to higher trophic levels (3–5). Diatoms can also be infected by smaller eukaryotic parasites and parasitoids (parasites that kill their hosts) as demonstrated by sporadic observations of "epidemics" within some diatom populations over the past several decades (6–8). While amplicon sequencing surveys have revealed a high diversity of putative eukaryotic parasites and many potential interactions with diverse diatoms (9, 10), the time, cost, and technical expertise required to make direct observations of diatom parasitism have thus far limited understanding of its prevalence, spatiotemporal dynamics, and oceanographic forcings (8).

One known example of diatom parasitism involves infection of the diatom, *Guinardia delicatula*, by the Cercozoan nanoflagellate, *Cryothecomonas aestivalis* (11). In this system, the parasitoid penetrates the host's cell wall and subsequently consumes the host protoplast from within, resulting in host mortality. After consuming its host, the enlarged *C. aestivalis* cell divides several (usually 3, occasionally more) times within the frustule to create swarmers that reenter the free-living plankton community, often leaving fecal pellets behind in the host frustule (11). *C. aestivalis* is known to infect *G. delicatula* in the North and Wadden Seas (7). More recently, a study at the nearshore Martha's Vineyard Coastal Observatory (MVCO) found recurrent, widespread infection of *G. delicatula* by *C. aestivalis* in a 7-y time series of manually annotated plankton images collected by Imaging FlowCytobot (IFCB) (12). In that study, infection by *C. aestivalis* was often found to regulate the magnitude of *G. delicatula* accumulation. When temperatures fell below 4 °C, however, infection was suppressed and accumulation of the host was frequently observed (12). This study combined with recent advances in automated image classification (13–15) demonstrates a path to develop automated, high-throughput approaches to quantify diatom parasitism on

Author contributions: D.C., E.E.P., and H.M.S. designed research; D.C., E.E.P., E.T.C., B.L.F.S., R.J.G., W.G.Z. and H.M.S. performed research; J.F., S.B., and H.M.S. contributed new reagents/analytic tools; D.C., E.E.P., E.T.C., and W.G.Z. analyzed data; and D.C. and H.M.S. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: dylan.catlett@whoi.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2303356120/-/DCSupplemental.

Published July 3, 2023.

unprecedented spatiotemporal scales, and in turn fill a major gap in our knowledge of the prevalence, ecological significance, and oceanographic determinants of infection.

Here, we develop and apply an automated, machine learning–based classifier to >1 billion IFCB images gathered from >300,000 samples from the Northeast U.S. Shelf (NES) to quantify *G. delicatula* abundance and infection by *C. aestivalis*. The classifier is trained, optimized, and validated with a large set of manually annotated *G. delicatula* images including uninfected and infected chains. We apply the classifier to quantify *G. delicatula* abundance and infection prevalence with daily resolution across an in situ, ~15-y time series at MVCO and with ~10 km resolution across a spatial domain extending from the northern Gulf of Maine to the southern Mid-Atlantic Bight sampled during 23 cruises conducted across 9 y. Our results support a previous study's findings of recurrent, significant levels of infection of *G. delicatula* and a suppression of infection at temperatures <4 °C at the nearshore MVCO. Use of the automated classifier enabled discovery of continued suppression of infection for at least 60 d following these cold temperatures and observations of comparable infection dynamics across a spatial domain spanning from Cape Hatteras to Nova Scotia. Integrating genetic observations with the imaging results suggested high niche diversity within the *C. aestivalis* species complex and that cold temperatures result in local extinction of the strain parasitizing *G. delicatula*. A long-term reduction in the spatiotemporal extent of waters with temperature <4 °C indicates that continued warming on the NES may reshape *G. delicatula* abundance and infection dynamics. In addition to elucidating the large-scale biogeography of *C. aestivalis* infection of *G. delicatula* on the NES, this study demonstrates the potential for automated plankton imaging coupled with artificial intelligence to advance knowledge of the understudied dynamics of phytoplankton–parasite interactions.

Results

Spatiotemporal Dynamics of *G. delicatula* Abundance and Infection on the NES. *G. delicatula* was observed at relatively high abundances (>5 chains mL^{-1}) throughout the NES and was often found at extremely high abundance (>50 chains mL^{-1}) at the nearshore MVCO time series site (Fig. 1). A crossshore gradient in *G. delicatula* abundances was evident across most of the NES, with high abundances close to shore declining to near-zero abundances beyond the shelf break (Fig. 1*A*). Low abundances were also observed in deep waters of the Gulf of Maine. *G. delicatula* abundances were highest in the northern Mid-Atlantic Bight and on Georges Bank, with relatively high abundances also seen on the Scotian Shelf. *G. delicatula* infection prevalence (Fig. 1*B*) was low (<5%) across most of the NES, but the spatial distribution showed a similar pattern to that of total *G. delicatula* abundance on the northern Mid-Atlantic Bight and Georges Bank. Areas associated with relatively high infection in these regions generally mirrored those associated with high abundance. Notably, infection was limited in nearshore waters of the Scotian Shelf despite relatively high *G. delicatula* abundances. Low *G. delicatula* abundance in offshore waters of the NES and in most of the Gulf of Maine drove high uncertainty in infection prevalence. Mean sea surface temperatures (SST) (Fig. 1*C*) across the NES were typically >20 °C offshore of the shelf break and

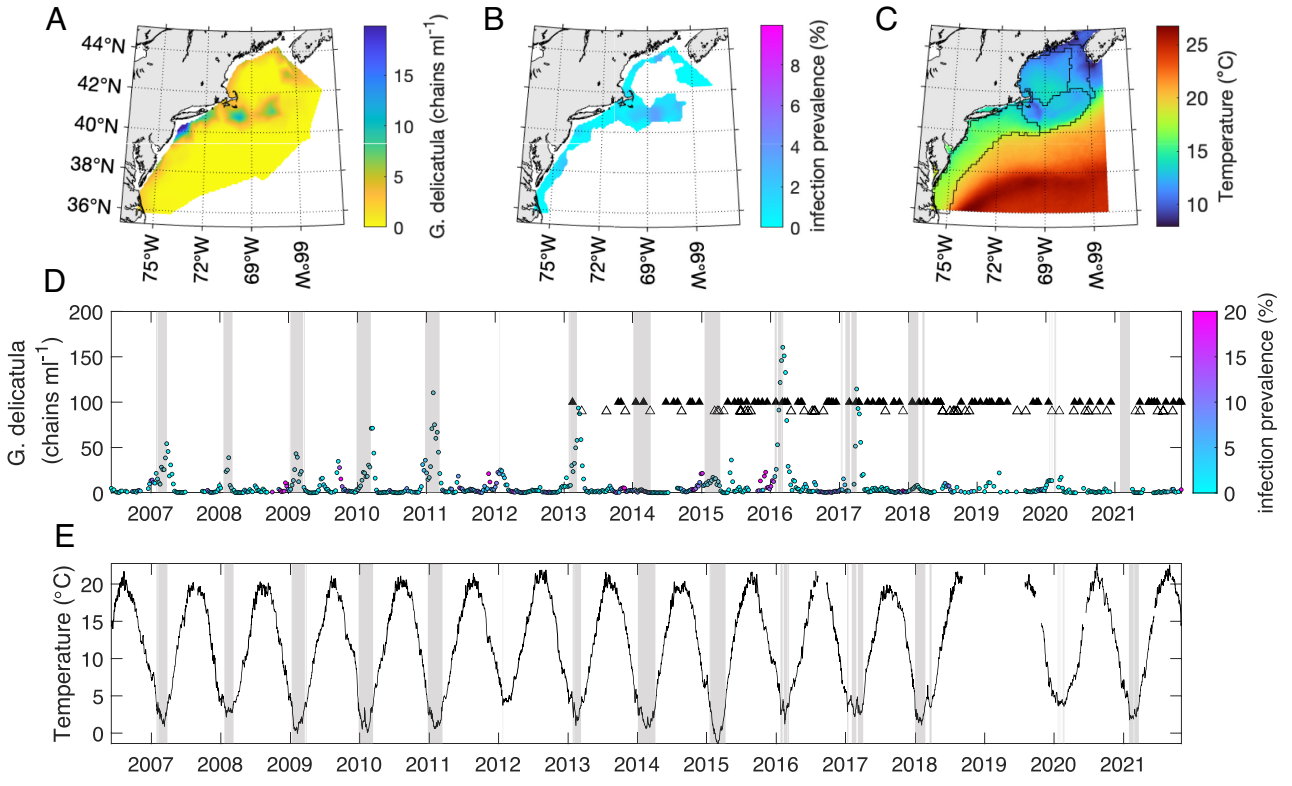

Fig. 1. Spatiotemporal dynamics of *G. delicatula* abundance and parasitoid infection. Maps show composites of (*A*) total *G. delicatula* abundance and (*B*) infection prevalence and (*C*) mean satellite sea surface temperature during the period over which these cruises took place. The color bar in (*A*) is truncated at 20 chains mL^{-1}. Black lines in (*C*) show the boundaries of "Ecological Production Units" previously defined for the Northeast U.S. Shelf (16). Infection prevalence is not shown where host abundance is <1 chain mL^{-1} due to high uncertainty. Time series observations are daily (*D*) aggregate *G. delicatula* abundance (dots) and infection prevalence (dot color) and (*E*) mean in-situ water temperature at MVCO. The *y* axis and colorbar in (*D*) are truncated at 200 chains mL^{-1} and 20%, respectively. Pink triangles in (*D*) from 2013 to 2021 indicate where at least one *C. aestivalis* ASV was (filled) or was not (open) detected in amplicon sequencing observations. Gray shading indicates periods where water temperature was <4 °C.

in the southern portion of the Mid-Atlantic Bight, 15 to 20 °C in the northern Mid-Atlantic Bight, and 10 °C or colder in the northern Gulf of Maine and Scotian Shelf. Infection prevalence appeared to be lowest in the cooler northern Gulf of Maine and Scotian Shelf, while high aggregate *G. delicatula* abundances were found across a broad range of mean temperatures except >20 °C.

Daily IFCB time series of *G. delicatula* abundance and infection prevalence alongside daily mean water temperature at MVCO provide a detailed view of temporal dynamics at a nearshore site on the NES, with intermittent amplicon sequencing observations providing additional insight into the dynamics of the parasitoid at this site (Fig. 1 *D* and *E*). *G. delicatula* abundance and infection prevalence were often higher at MVCO than across the broader NES, regularly reaching abundances >50 chains mL^{-1} and infection prevalence >20%. High abundances were most often observed in the winter during or immediately following periods where water temperature was <4 °C and infection prevalence was negligible, in agreement with previous observations (12). *G. delicatula* also accumulated at other times of year when temperatures were warmer, but during these events, abundances rarely exceeded 50 chains mL^{-1}. Parasitoid infection was most prevalent in the fall, with infection prevalence values of 20% or higher regularly observed prior to the seasonal onset of temperatures <4 °C. Infection prevalence was typically low during the spring and early summer, even when *G. delicatula* reached moderately high abundances and temperature exceeded 4 °C. Interestingly, at least some of the five *C. aestivalis* amplicon sequence variants (ASVs, a genetic proxy for strains or species) were detected at MVCO throughout the year, including during periods cooler than 4 °C and in the spring and summer when *G. delicatula* infection prevalence was negligible.

Observations show seasonal variations in the distributions of *G. delicatula* abundance and infection prevalence and in SST across the NES (Fig. 2). While only one cruise with limited coverage of the Gulf of Maine, Scotian Shelf, and southern Mid-Atlantic Bight was conducted during the winter, *G. delicatula* abundances were elevated (>20 chains mL^{-1}) and infection prevalence was high compared to other seasons across much of the NES. Mean winter-time SSTs on the NES were generally <10 °C and were <4 °C in many

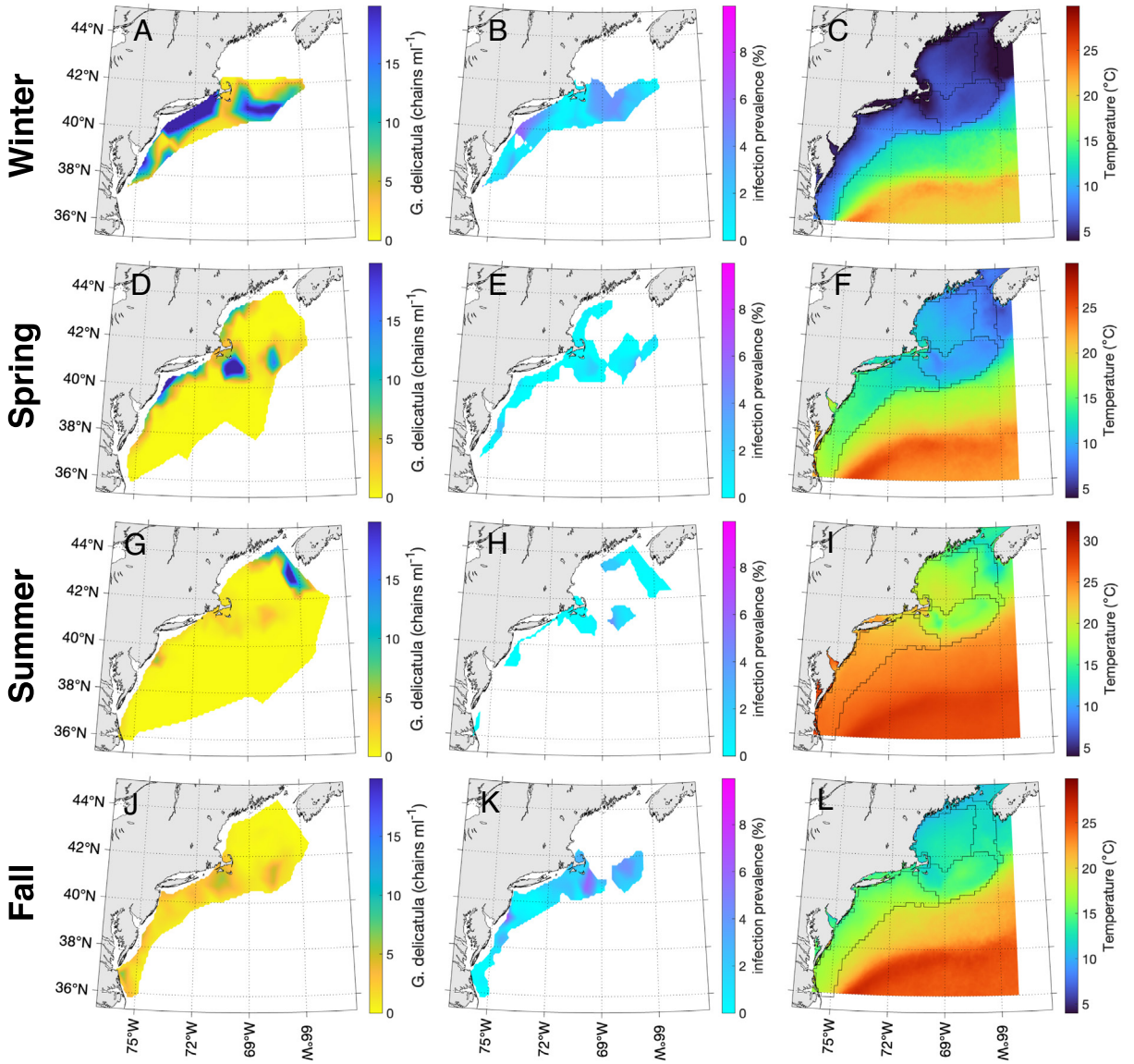

Fig. 2. Composite maps of total *G. delicatula* abundance and infection prevalence and mean satellite sea surface temperature during (*A-C*) winter, (*D-F*) spring, (*G-I*) summer, and (*J-L*) fall from 2013 to 2021. Infection prevalence is not shown where host abundance is <1 chain mL^{-1} due to high uncertainty. Abundance colorbars are truncated at 20 chains mL^{-1}, infection prevalence colorbars are truncated at 10%, and temperature colorbars are truncated at 4 °C.

nearshore locations. *G. delicatula* abundance remained high across much of the Mid-Atlantic Bight and Georges Bank during the spring, with a notable nearshore band of elevated abundance also observed in the Gulf of Maine. Where infection prevalence was measurable during the spring, it was generally lower than that in winter despite mean SSTs ≥~10 °C. *G. delicatula* abundances were low (<5 chains mL^{-1}) across much of the NES during the summer, and infection prevalence was generally not quantifiable when seasonal mean SST was highest across most of the domain. The exception was at the northern edge of the domain including the Scotian Shelf and a small portion of the northern Gulf of Maine where high *G. delicatula* abundances coincided with negligible infection and relatively cool SSTs. In the fall, *G. delicatula* was typically not detected throughout the Gulf of Maine and Scotian Shelf, while intermediate abundances (~5 chains mL^{-1}) and SSTs associated with relatively high (>5%) infection prevalence were observed in several locations across the Mid-Atlantic Bight and Georges Bank.

Spatial and Temporal Gradients in *G. delicatula* Abundance and Infection. Distinct annual cycles were observed in *G. delicatula* abundance and infection prevalence both at MVCO and across the broader NES (Fig. 3 *A* and *B*). Monthly aggregate *G. delicatula* abundances were highest in the winter and spring, while infection prevalence was highest during the fall. From October to December at MVCO, aggregate *G. delicatula* abundance was approximately 5 chains mL^{-1} or less, coinciding with high aggregate infection prevalence of ~10 to 12%. The monthly aggregate abundances increased to ~10 chains mL^{-1} in January until reaching an annual maximum of >30 chains mL^{-1} in March. This increase in monthly aggregate abundance coincided with a decline in aggregate infection prevalence to ~4% in January and an annual minimum of <1% in March. Comparable patterns were observed across the broader NES, although monthly aggregate abundance and infection prevalence values were lower and more uncertain due to less intensive sampling and lower host abundances.

A pronounced crossshore decline in *G. delicatula* abundance was observed across the NES (Fig. 3*C*), with aggregate abundances of ~7 chains mL^{-1} within 50 km of the coastline declining to 2.30 chains mL^{-1} between 50 and 100 km from the coast and <2 chains mL^{-1} beyond 100 km from the coast. Consistent crossshore gradients in infection prevalence were not observed, although values were generally higher offshore. Across the four NES Ecological Production Units (16, 17), relatively high, aggregate *G. delicatula* abundances (>2.5 chains mL^{-1}) were observed in the Mid-Atlantic Bight and on Georges Bank and the Scotian Shelf, with lower aggregate abundances (<1.5 chains mL^{-1}) found in the Gulf of Maine. Aggregate infection prevalence ranged from 0.7 to 1.3% in the Mid-Atlantic Bight, Georges Bank, and Gulf of Maine regions but was reduced (<0.5%) on the Scotian Shelf.

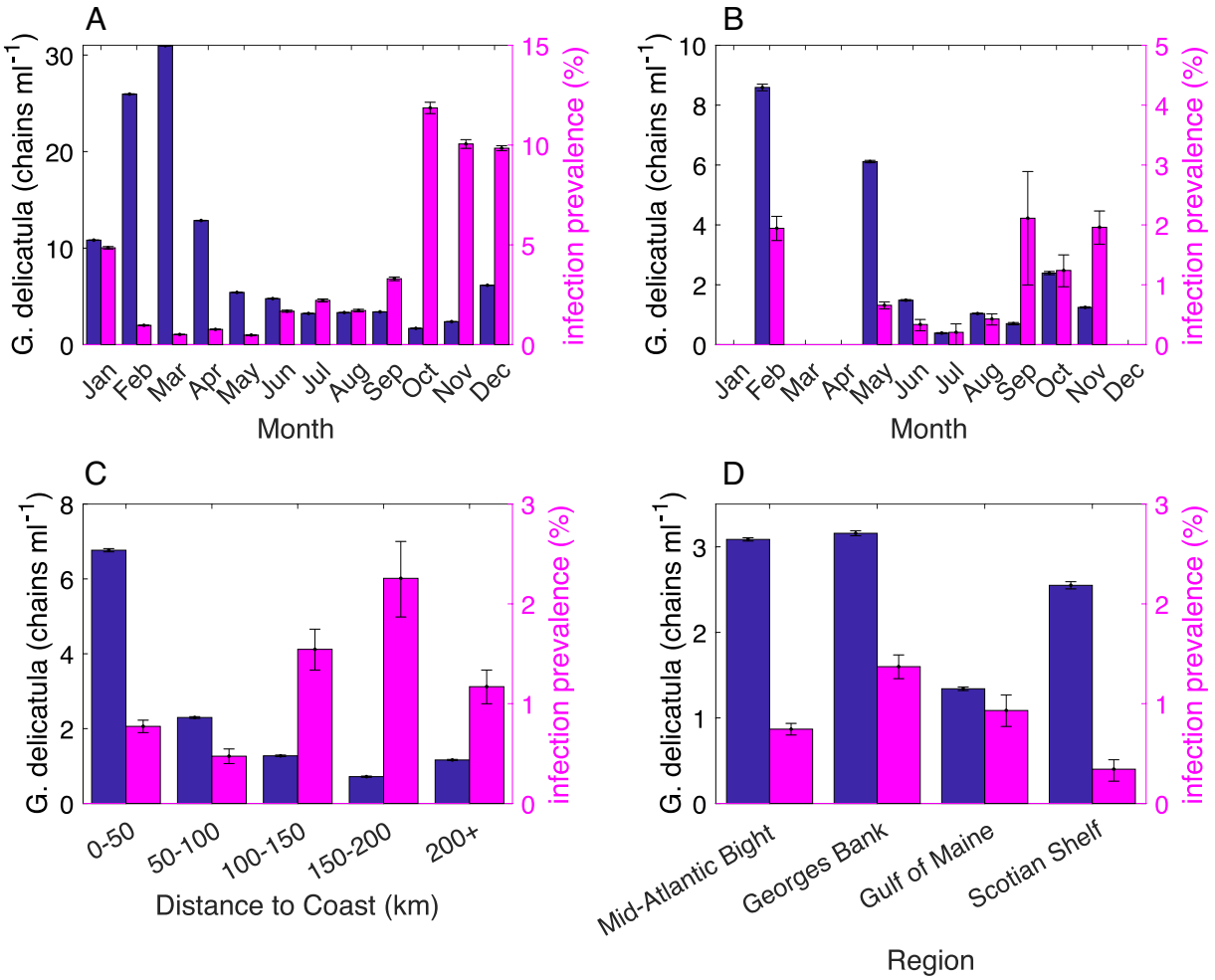

Fig. 3. Spatiotemporal gradients in *G. delicatula* abundance (navy) and infection prevalence (pink). Shown are (*A* and *B*) monthly aggregate abundance and infection prevalence (*A*) at the nearshore Martha's Vineyard Coastal Observatory (MVCO) time series and (*B*) across the broader Northeast U.S. Shelf, and (*C*) crossshore and (*D*) interregion gradients. (*B*–*D*) do not consider observations at MVCO. Error bars show 95% CIs for aggregate concentrations assuming that counts are drawn from a Poisson distribution. Note that scales vary across different panels.

Role of Temperature in Governing *G. delicatula* Infection. Previous observations (12) and our results suggest that water temperatures <4 °C provide *G. delicatula* with a refuge from parasitoid infection. The more extensive results we report here confirm the critical role of temperature in governing parasitoid infection of *G. delicatula* at MVCO (Fig. 4*A*) and across the broader NES (Fig. 4*B*). At MVCO, high *G. delicatula* abundance coincident with low infection prevalence (typically <2%) was observed frequently at temperatures <4 °C as observed previously via manual image classification (12). High *G. delicatula* abundances were also regularly observed when temperatures exceeded 4 °C but were often associated with nonnegligible (>5%), and at times high (>20%), infection prevalence. Observations across the broader NES made possible by the automated classifier supported this pattern, with negligible infection prevalence found in all samples collected in waters colder than 4 °C except one (associated with high uncertainty). A small number (22 of 4,219) of daily observations at MVCO had nonnegligible infection prevalence (>5%) coincident with *G. delicatula* abundances >3 chains mL^{-1} at temperatures <4 °C. Most (18 of 22) of these daily observations were due to erroneous automated image classification, while the others were a result of high-resolution sampling and uncertainty in temperature observations (*SI Appendix*, *Supporting Text*).

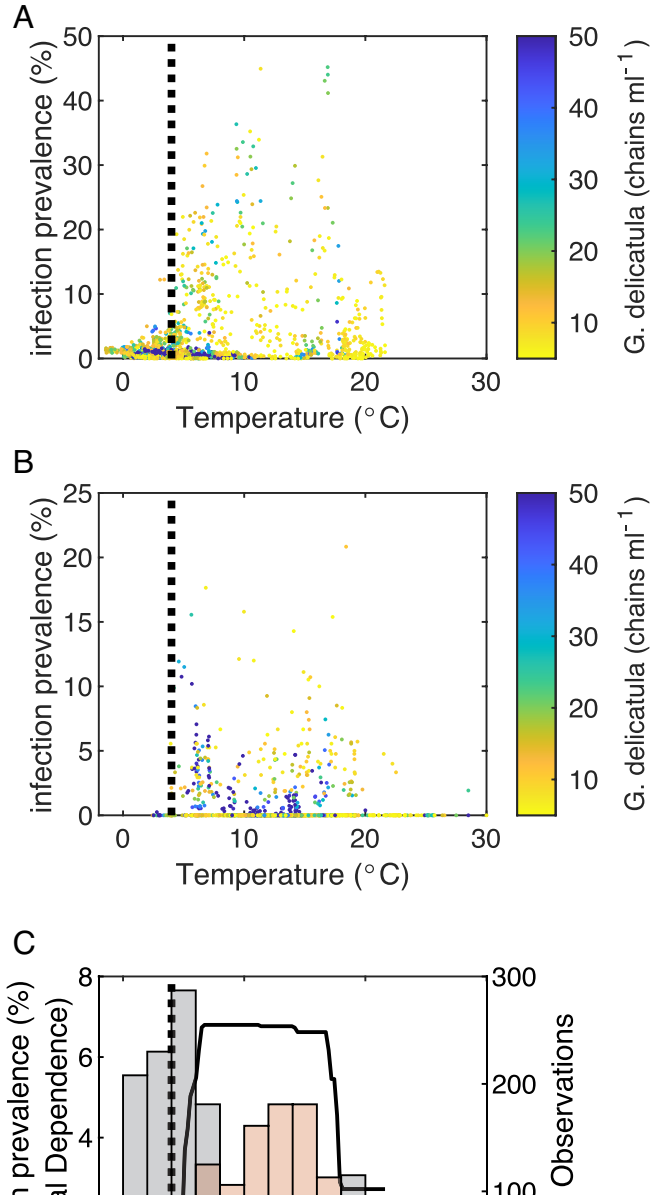

Fig. 4. Role of temperature in governing *G. delicatula* parasitoid infection dynamics (*A*) at MVCO and (*B*) across the broader NES. (*C*) shows the partial dependence of infection prevalence on temperature determined from regression tree analysis (left axis) and the number of observations in each temperature bin (right axis) at MVCO (black) and across the broader NES (red). In all panels, vertical dashed lines indicate 4 °C and observations are only considered where *G. delicatula* abundance is >5 chains mL^{-1}. Observations at MVCO are daily aggregate quantities. The colorbar is truncated at 50 chains mL^{-1} to facilitate visualization of abundances.

A bootstrap-aggregated ensemble of 100 regression tree models fit to predict infection prevalence from temperature showed the largest change in the partial dependence of infection prevalence at temperature thresholds of 4.6 and 4.1 °C in the MVCO and shipboard data, respectively (Fig. 4*C*). In both datasets, infection prevalence partial dependence increased substantially at these thresholds. At MVCO, the partial dependence of infection prevalence was elevated at temperature values ranging from 4.6 to 17.6 °C and showed a sharp decline when temperatures were >17.6 °C. However, the number of daily observations of elevated *G. delicatula* abundance was also lower around this upper threshold relative to the 4.6 °C threshold. In shipboard observations, the partial dependence of infection declined and remained low at temperatures >7.5 °C.

The high-resolution time series and well-resolved annual cycles in *G. delicatula* abundance and infection enabled by the automated classifier (Figs. 2–4) suggest that parasitoid infection of *G. delicatula* continues to be suppressed for some time after temperatures have warmed above 4 °C in spring. We investigated this further by computing the frequency of days at MVCO with *G. delicatula* abundance >5 chains mL^{-1}, coinciding with and without infection prevalence >5% as a function of the number of days since the last observation of mean daily temperature <4 °C in each year (which we refer to hereafter as the end of the "cold snap"; Fig. 5). Across 14 y at MVCO with adequate data availability (Fig. 1 *D* and *E*), *G. delicatula* was found at high abundance frequently (>300 total days) within 60 d of the end of the cold snap, but <2% of these days coincided with infection prevalence >5% (Fig. 5*A*). High *G. delicatula* abundances were found in conjunction with high infection prevalence within 30 d of the cold snap only in 2012 when the preceding cold snap lasted only 3 d. Observations of infection events were also rare within 60 d of the end of the cold snap, only occurring in 2012 and 2013 and accounting for a cumulative total of 6 of the 88 d where high *G. delicatula* abundances were observed (Fig. 5*B*). Variable patterns of detection of the five *C. aestivalis* ASVs at MVCO, including frequent detection of two ASVs during and following the cold snap, suggest that some strains of *C. aestivalis* survive the cold snap (Fig. 5*C*). Interestingly, however, the most abundant *C. aestivalis* ASV at MVCO was only detected from June to December (Fig. 5*C*) and was never detected during or within 70 d of the termination of the cold snap.

Long-Term Trend in the Duration and Areal Extent of Cold Snaps on the NES. A long-term warming trend of 0.37 °C $decade^{-1}$ was recently reported on the NES (18). Our results suggest that this warming trend could have significant impacts on the dynamics of *G. delicatula* abundance and infection by *C. aestivalis* if it

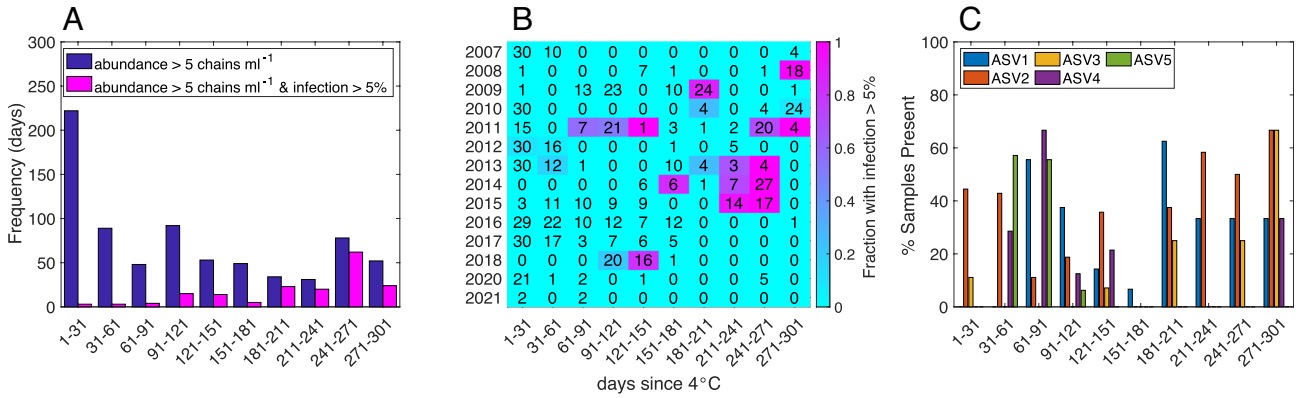

Fig. 5. Continued suppression of infection following cold periods. (*A*) shows the number of days with aggregate *G. delicatula* abundance >5 chains mL^{-1} (navy) coincident with infection prevalence >5% (pink) relative to the number of days since daily average temperatures were <4 °C across all years at Martha's Vineyard Coastal Observatory (MVCO). (*B*) shows these data for individual years, with black numbers showing the number of days with aggregate *G. delicatula* abundance >5 chains mL^{-1} and tile color indicating the fraction of those days with infection prevalence >5%. (*C*) shows relative frequencies of detection of the five *C. aestivalis* ASVs detected in amplicon sequencing observations relative to the number of days for detection of each ASV following an observation of water temperature <4 °C across all years at MVCO. *C. aestivalis* ASVs are numbered according to their rank order cumulative relative abundance across all available samples, with ASV1 being the most abundant.

contributes to a decline in the duration and areal extent of cold snaps. More detailed analysis of the satellite SST record from 1982 to 2022 showed that the areal extent of cold snaps on the NES follows a clear seasonal cycle, with the maximum areal extent observed during the winter as expected (Fig. 6*A*). However, over this time period, the spatiotemporal extent of cold snaps has declined at an average rate of -2.1×10^4 d km y^{-1}, indicating a likely reduction in the extent of *G. delicatula*'s thermal refuge from parasitism in this region (Fig. 6*B*). Extreme realizations of short-lived and spatially restricted cold snaps were also more frequent in the final decade of the 40-y satellite record, with the 5 shortest and/or spatially restricted cold snaps observed in 2012, 2016, 2020, 2021, and 2022 (Fig. 6*B*).

Discussion

Summary. We used automated, in situ imaging and a convolutional neural network (CNN) image classifier to quantify, characterize the spatiotemporal distributions and gradients in, and determine the role of temperature in controlling *G. delicatula* abundance and parasitoid infection on the NES. We found that *G. delicatula* reached high abundance (>20 chains mL^{-1}) throughout the NES. The highest *G. delicatula* abundances were observed in the winter and spring in the nearshore Mid-Atlantic Bight and on Georges Bank, with high abundances also found during the summer on the Scotian Shelf. Parasitoid infection of *G. delicatula* was most prevalent in the fall in the Mid-Atlantic Bight and on Georges Bank but became negligible when temperatures fell below 4 °C. Here, we discuss the physical and biological controls on *G. delicatula* abundance and infection dynamics and their implications with respect to ongoing warming of the NES surface ocean. We conclude with a discussion of the current limitations and future potential for automated imaging and machine learning classification to advance understanding of phytoplankton–parasite interactions.

Gradients in and Controls on *G. delicatula* Abundance and Infection Dynamics on the NES. Since *G. delicatula* is one of the biomass-dominant diatoms on the NES (12), understanding its abundance dynamics and the fates of its production under variable oceanographic conditions is an important step in predicting the response of the NES ecosystem to anthropogenic climate forcing. Our results revealed two prominent gradients in *G. delicatula* abundances. First, a prominent crossshore gradient was found, with the highest abundances consistently observed close to shore (Figs. 1–3). This gradient in *G. delicatula* abundances follows the cross-shelf decline in phytoplankton and diatom biomass previously observed across the NES (19–22). These gradients are typically attributed to increased nutrient loading in the nearshore NES due to terrestrial and estuarine inputs (21, 22), which likely contributes to the cross-shore gradient in *G. delicatula* abundances observed here. Conversely, cross-shore gradients in *G. delicatula* infection prevalence were not resolved here due to high uncertainty in offshore estimates. Nonetheless, persistently higher aggregate *G. delicatula* abundances and infection prevalence found at the nearshore MVCO relative to the cruise observations (Figs. 1–3) suggest that infection prevalence may be elevated very close to shore where *G. delicatula* abundances tend to be especially high. Studies have shown that infection in some freshwater diatom–parasite systems is host density dependent with "epidemics" only observed at relatively high host densities (23), which may explain

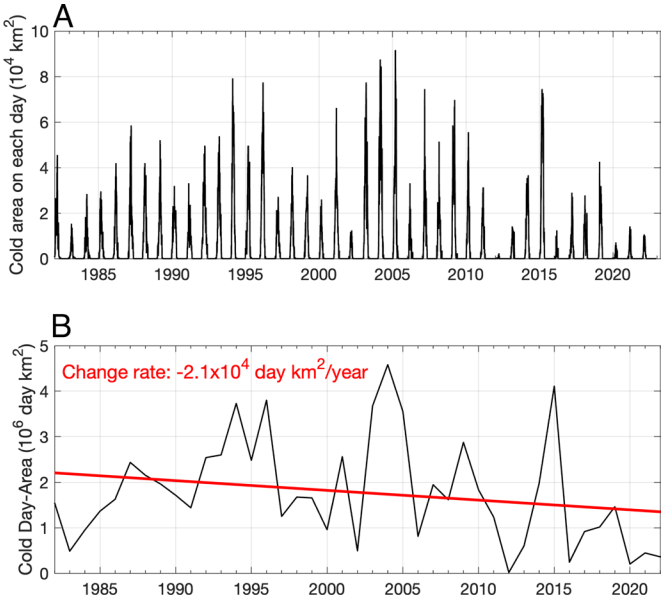

Fig. 6. Long-term trend in the spatiotemporal extent of cold periods on the Northeast U.S. Shelf (NES). Shown are (*A*) a daily time series of the areal extent of waters with sea surface temperature (SST) <4 °C on the NES and (*B*) its annual integral. The red line in (*B*) indicates the long-term linear trend determined by linear regression.

the differences in infection prevalence between the nearshore MVCO and the broader NES.

A second prominent gradient in *G. delicatula* abundances and infection revealed here is the annual cycle (Figs. 1–3). The largest bloom events typically occurred during or immediately following periods with water temperatures colder than 4 °C, while parasitoid infection was negligible (Figs. 1, 4, and 5), suggesting that temperature may indirectly control *G. delicatula* abundance through its impacts on infection (12). Observations at MVCO further suggest that *G. delicatula* infection remains rare for 60+ d following the cessation of cold temperatures (Figs. 1, 3, and 5). Elevated *G. delicatula* infection prevalence during the fall and winter followed by minimal infection in the spring across much of the Mid-Atlantic Bight and Georges Bank (Fig. 2) suggests that this pattern holds across much of the NES. Further, from July to September, *G. delicatula* was highly abundant while infection was negligible in the northern-most portion of the domain (Fig. 2) where climatological mean SST remains below 4 °C until May (24). Interestingly, the few published observations of *G. delicatula* infection by *C. aestivalis* in other systems (the North and Wadden Seas) were also made during the late summer and fall (7, 11). Together, these observations suggest that the annual cycle in *G. delicatula* abundance and infection prevalence may vary spatially across the NES, with northern, colder areas providing longer *G. delicatula* "accumulation windows" by suppressing parasitoid infection. Regression tree analysis suggested that infection in this system may also be suppressed at temperatures >18 °C, but the relatively few observations of elevated *G. delicatula* abundance at these temperatures limit interpretation of this upper bound in the thermal range of infection.

Elucidating the biological mechanisms driving the response of *G. delicatula* infection to temperature will require experimental verification, but our results aid in the development of hypotheses to explain this phenomenon. Some of the five *C. aestivalis* ASVs detected in amplicon sequencing observations at MVCO were present throughout the year, including when temperatures were <4 °C and in the spring and summer following these periods (Figs. 1*D* and 5*C*). The presence of some *C. aestivalis* strains during prolonged periods of negligible *G. delicatula* infection shows that some representatives of the *C. aestivalis* species complex likely survive at temperatures <4 °C. Interestingly, the most abundant *C. aestivalis* ASV was not detected during or within 70 d following cold periods, suggesting that microdiversity within the *C. aestivalis* species complex may drive extreme host specificity where only some strains (apparently one at MVCO) infect *G. delicatula*. Differences in host and parasite thermal tolerances have been shown to drive similar responses of diatom abundance and infection to temperature in freshwater diatom–parasite systems (23, 25). We thus hypothesize that temperature-driven local extinction events of a single strain of *C. aestivalis* explains the suppression of *G. delicatula* infection during cold snaps, and that infection is only restored after physical processes "reseed" this strain of *C. aestivalis*. Unfortunately, the limited accuracy of DADA2 in inferring ASVs from nonoverlapping paired reads (26) prohibits definitive statements regarding the importance of *C. aestivalis* microdiversity relative to other potential explanations, such as life cycle transitions or physiological adjustments of *C. aestivalis* to cold temperatures, in explaining the *G. delicatula* infection dynamics observed here.

Our results have important implications for understanding future variability in the NES ecosystem as it is one of the fastest-warming marine ecosystems on Earth (18, 27, 28). The long-term decline in the duration and areal extent of waters with temperatures <4 °C and our observations of *G. delicatula* abundance and infection (Figs. 4–6) suggest that dramatic changes in the magnitude and phenology of *G. delicatula* accumulation may already be underway on the NES. If these long-term trends continue, impacts might include reduced or a varied composition of phytoplankton biomass available for consumption by larger micro- and mesozooplankton [which may have cascading impacts on the NES food web; (29)] and/or decreased export of diatom biomass and aggregates from the surface ocean to depth (30, 31). These implications point toward an urgent need to develop a predictive understanding of the forcings and impacts of diatom parasitism in economically important coastal ocean ecosystems subject to increasing anthropogenic influence like the NES.

Potential and Limitations of Automated Imaging and Classification to Quantify Phytoplankton Parasitism. Microbial interactions shape planktonic food webs and elemental cycles throughout the worlds' oceans (32, 33), but to date, these interactions have been difficult to quantify across large spatiotemporal scales due to technical limitations. Parasitism is an understudied planktonic interaction given the proliferation of evidence suggesting that parasites are highly abundant members of marine plankton communities (9, 10). Since diatoms exert disproportionate influence on marine ecosystems, foundational knowledge of the spatiotemporal dynamics and ecology of diatom–parasite interactions is needed to better understand and predict the flow(s) of organic matter through marine food webs (8). This study demonstrates how high-throughput, automated image analysis can be used to quantify diatom parasitism. Here, we discuss the potential and limitations of this approach to advance our understanding of diatom–parasite interactions in the marine environment, and of phytoplankton–parasite interactions in aquatic systems more broadly.

Currently, plankton imaging provides sufficient resolution to identify many microplankton genera and some species. Limited image resolution in combination with low morphological diversity in smaller sized plankton largely precludes taxonomic identification and will likely prevent image-based observation of both free-living parasites and smaller hosts for some time. Our identification of the parasitoid of *G. delicatula* as *C. aestivalis*, and of infected *G. delicatula* images, relied on close agreement of the imaged infection cycle stages with previous, detailed observations of the *C. aestivalis* infection cycle in culture (11, 12). Some uncertainty remains as to whether all infection events identified here were carried out by *C. aestivalis* or another parasitoid with morphologically similar infection dynamics. These caveats demonstrate important considerations for identifying and quantifying parasitic infection in other diatoms and microplankton with automated imaging. Morphological features of infection must be known and resolvable in plankton images to identify infected hosts. These morphological characteristics can be leveraged in cases where the parasitoid is not directly observed, as in morphological responses of *Alexandrium fundyense* to *Amoebophyra sp.* infections (34). In cases where infection cycles or host morphological responses to infection are unknown, recent work (32) suggests that supplementing images with metagenetic, genomic, or transcriptomic observations to support the identification of likely interactions of known or suspected parasites with new hosts may allow for application of these approaches to currently undocumented microplankton–parasite interactions.

Our results highlight several areas where technical advances in plankton imaging and image analysis will facilitate efforts to improve quantitation of diatom and microplankton parasitism. First, reducing uncertainties in estimates of host abundance and infection is possible through continued development of both the instrumentation and machine learning approaches employed here. Higher-throughput plankton imaging (35) will reduce sampling uncertainty. Similarly, improvements in automated image classifiers will reduce classification uncertainty. Both improvements will

reduce the degree of sample aggregation needed to resolve spatiotemporal gradients and increase the resolution of small-scale infection dynamics. Further, improved image classifiers may allow for resolution of additional stages of infection: Preliminary classifiers tested in the present study were unable to consistently distinguish *G. delicatula* chains with attached external nanoflagellates (potentially parasitoids) from other *G. delicatula* morphologies. The automated classifier employed here also required an extensive collection of human-annotated images to ensure consistency and accuracy in both infected and noninfected *G. delicatula* classifications across the billions of images analyzed here. Improvements in machine learning classification may also reduce the required number of human annotations to carry out large-scale investigations of population dynamics. While our optimized "classify and count" approach was highly effective for the applications used here, other studies suggest that automated quantifiers may improve estimates of plankton abundance from image data (36, 37). These methods should be more rigorously evaluated for quantification of diatom–parasite interactions. Finally, well-validated segmentation algorithms to quantify both infected and noninfected diatom cell abundances from IFCB or other image data (38) are not currently available, and our study thus focuses on the combined chain and cell abundance and infection prevalence of *G. delicatula*. These and other technical developments will improve resolution of the dynamics and unlock powerful studies of the ecology of microplankton– and diatom–parasite interactions.

Materials and Methods

Imaging-In-Flow Cytometry. IFCB has been deployed in situ at ~4 m depth at the MVCO offshore tower located in ~15 m water depth off the southern coast of Martha's Vineyard, MA, USA, since 2006 (12, 39). IFCB collects discrete seawater samples every ~25 min and images particles larger than ~6 μm. Manually classified IFCB images are available from regular biweekly monitoring of IFCB data as well as a previous study that identified *G. delicatula* parasitism in IFCB images from MVCO (12). IFCB has also sampled quasi-continuously from the underway flow-through system on 24 (23 of which are considered here; see *SI Appendix, Supporting Text*), approximately quarterly cruises (conducted by the National Oceanic and Atmospheric Administration EcoMon program, e.g., ref. 40) that surveyed the NES since 2013. Winter is undersampled with only one cruise available, but all other seasons included seven or more cruises. Additional details are provided in *SI Appendix, Supporting Text*.

Amplicon Sequencing. Amplicon sequencing observations from MVCO are considered to determine the presence or absence of *C. aestivalis* in the water column independently from observations of *G. delicatula* infection (free-living *C. aestivalis* cannot be unambiguously identified in IFCB images). A total of 135 samples were collected for amplicon sequencing of the V4 hypervariable region of the 18S rRNA gene from February 2013 to December 2021. 18S rRNA gene amplicons were generated with the 574*f and 1132r primers from ref. 41 (CGGTAAYTCCAGCTCYV; CCGTCAATTHCTTYAART) and sequenced on an Illumina MiSeq (see *SI Appendix, Supporting Text* for complete methods). Demultiplexed sequence data were processed using the DADA2 method (26) with some modifications to accommodate nonoverlapping paired reads (*SI Appendix, Supporting Text*). BLAST analysis (42) showed that 5 ASVs were perfect matches to at least one *C. aestivalis* reference sequence included in the Protistan Ribosomal Reference database v4.14.0 (43); these ASVs are considered here.

In Situ and Remotely Sensed Temperature Observations. Water temperature observations at MVCO are derived from a SeaBird Electronics MicroCAT CTD deployed alongside the IFCB at the offshore tower. Where water temperature observations are not available, observations from another MicroCAT CTD deployed at ~12 m depth at an undersea node ~1.5 km from the IFCB deployment site are used to fill gaps in the offshore tower's temperature time series. Across the broader NES, a combination of underway water temperature and satellite-measured SST observations is considered. Underway observations are considered for estimates of water temperature concurrent with IFCB observations (Fig. 4), while satellite SST is used in composite maps (Figs. 1 and 2) and long-term trend analyses (Fig. 6). *SI Appendix, Supporting Text* provides detailed analysis methods for these data.

Automated Image Classification. A CNN image classifier with Inception v3 architecture (13) was used for automated classification of IFCB images. A training set of at least 20 and at most 2,000 manually annotated images from each of 155 classes was used for training and initial validation of the classifier. The classifier was initialized with pretrained weights from ImageNet (44) and fine-tuned with the NES IFCB training set (97,026 images, 155 classes, 80 to 20 split for training and validation). The classifier performed well in initial validations based on the 20% hold-out set with an F1 statistic (harmonic mean of precision and recall) across all classes of 0.91.

Two classes of *G. delicatula* were separated by the CNN classifier: *G. delicatula* and infected *G. delicatula* (*SI Appendix*, Fig. S1). The "*G. delicatula*" class includes images of *G. delicatula* cells and chains exhibiting a range of morphological features, as well as some images depicting *G. delicatula* cells or chains with small flagellates (potentially parasitoids) apparently attached to the exterior of the frustule (*SI Appendix*, Fig. S1). *G. delicatula* images with externally attached nanoflagellates were included in the noninfected *G. delicatula* class rather than the infected *G. delicatula* class for two reasons: Host responses to external parasitoid attachment are unknown and could include successful defense against infection, and other nanoflagellates may interact with *G. delicatula*. The class "infected *G. delicatula*" includes images of *G. delicatula* cells or frustules that exhibit signs of current or recent parasitoid presence within the host frustule (*SI Appendix*, Fig. S1). The class-specific F1 statistics for automated classifications of *G. delicatula* and infected *G. delicatula* images using the "hold-out" subset of the classifier training set were 0.92 and 0.91, respectively.

Manually annotated images from MVCO independent from those considered in classifier training were used to optimize and evaluate classifier performance when applied to unique IFCB observations. Classifier optimization is described in *SI Appendix, Supporting Text* and sought to equate precision and recall statistics (and thus, false positive and negative classification errors) for *G. delicatula* and infected *G. delicatula* classifications. The F1 statistics obtained for *G. delicatula* and infected *G. delicatula* in the MVCO validation set after optimization were 0.92 and 0.78, respectively. Estimates of *G. delicatula* and infected *G. delicatula* abundances and *G. delicatula* infection prevalence based on human and automated classifiers were in close agreement with one another (*SI Appendix, Supporting Text* and Figs. S2 and S3).

The same optimized classifier was applied and additional classifier validation was performed for shipboard IFCB observations (*SI Appendix, Supporting Text*). In these data, automated *G. delicatula* classifications achieved an F1 value of 0.91 with precision and recall values of 0.92 and 0.90, respectively, while infected *G. delicatula* classifications had an F1 value of 0.50 and precision and recall values of 0.52 and 0.48, respectively. Following image classification, counts of *G. delicatula* and infected *G. delicatula* chains (it is currently not possible to enumerate individual cells in IFCB images of diatoms) were determined for discrete IFCB samples.

Spatiotemporal Distributions and Gradients. Maps were created with the m_map MATLAB toolbox (45). We used a spatial sample aggregation procedure to reduce uncertainty in *G. delicatula* abundance and infection estimates and smooth the data for inspection of large-scale spatial distributions (*SI Appendix, Supporting Text* and Figs. S4 and S5). Monthly aggregate abundances and infection prevalence were used to quantify seasonal gradients. To characterize crossshore gradients, the Climate Data Toolbox (46) was used to compute great circle distances from each IFCB sample to the nearest coastline. We characterize gradients across the four major "Ecological Production Units" on the NES (16, 17) after slightly modifying the boundaries used by ref. 16 to incorporate more of the available IFCB data (*SI Appendix*, Fig. S6). Spatiotemporal gradients across each of these dimensions are characterized by aggregating discrete IFCB samples and estimating uncertainty as described in *SI Appendix, Supporting Text*.

Regression Tree Analysis. A bootstrap-aggregated ensemble of 100 regression trees was fit to both the shipboard and daily aggregated MVCO datasets of concurrent *G. delicatula* infection prevalence and temperature observations. Our goal in this analysis was to objectively determine critical temperature thresholds that may result in dramatic changes in *G. delicatula*'s susceptibility to infection. For this reason, each regression tree was constrained to include only 2 splits. Concurrent temperature and infection prevalence values were only included in this analysis

where *G. delicatula* abundances were >5 chains mL^{-1} to prevent biasing results toward highly uncertain infection prevalence estimates. The partial dependences of infection prevalence on temperature in each dataset illustrate predicted values of infection prevalence from a given temperature value from the ensemble tree.

Data, Materials, and Software Availability. Data on automated and human classifier counts of infected and noninfected diatoms have been deposited in Environmental Data Initiative (10.6073/pasta/7c43f9e037325e70ee736b7782327dd8) (47).

ACKNOWLEDGMENTS. We acknowledge the captains and crews of the vessels used to collect the data presented here, National Oceanic and Atmospheric Administration's support of the survey cruises considered here, and the MVCO Operations Team for maintaining the facility. We are grateful to R. Olson and A. Shalapyonok for their dedicated contributions in maintaining the IFCB time series at MVCO. We also thank two anonymous peer reviewers and the editorial team at PNAS for their constructive and timely review of our manuscript. Amplicon sequencing was accomplished with the use of sequencing services at the Rhode Island Institutional Development Award Network of Biomedical Research Excellence from the National Institute of General Medical Sciences of the NIH under grant number P20GM103430 through the Centralized Research Core facility and/or the Molecular Informatics Core (RRID:SCR_017685). Financial support for this work was provided by the Simons Foundation (561126); the NSF (OCE-1434440; OCE-1851012), especially through the Northeast U.S. Shelf Long-Term Ecological Research (OCE-1655686); fellowship support (to H.M.S.) through the NOAA Cooperative Institute for the North Atlantic Region under Cooperative Agreement NA14OAR4320158; and fellowship support (to D.C.) by the NSF (OCE-2205596).

Author affiliations: [a]Department of Biology, Woods Hole Oceanographic Institution, Woods Hole, MA 02543; and [b]Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA 02543

1. J. H. Ryther, Photosynthesis and fish production in the Sea. *Science* **166**, 72–76 (1969).
2. P. Tréguer *et al.*, Influence of diatom diversity on the ocean biological carbon pump. *Nat. Geosci.* **11**, 27–37 (2018).
3. R. T. Barber, M. R. Hiscock, A rising tide lifts all phytoplankton: Growth response of other phytoplankton taxa in diatom-dominated blooms. *Global Biogeochem. Cycles* **20**, GB4S03 (2006).
4. V. Smetacek, Diatoms and the ocean carbon cycle. *Protist* **150**, 25–32 (1999).
5. E. B. Sherr, B. F. Sherr, Heterotrophic dinoflagellates: A significant component of microzooplankton biomass and major grazers of diatoms in the sea. *Mar. Ecol. Prog. Ser.* **352**, 187–197 (2007).
6. L. P. M. J. Wetsteyn, L. Peperzak, Field observations in the Oosterschelde (The Netherlands) on Coscinodiscus concinnus and Coscinodiscus granii (Bacillariophyceae) infected by the marine fungus Lagenisma coscinodisci (Oomycetes). *Hydrobiol. Bull.* **25**, 15–21 (1991).
7. U. Tillmann, K.-J. Hesse, A. Tillmann, Large-scale parasitic infection of diatoms in the Northfrisian Wadden Sea. *J. Sea Res.* **42**, 255–261 (1999).
8. B. Scholz *et al.*, Zoosporic parasites infecting marine diatoms–A black box that needs to be opened. *Fungal. Ecol.* **19**, 59–76 (2016).
9. G. Lima-Mendez *et al.*, Determinants of community structure in the global plankton interactome. *Science* **348**, 1262073 (2015).
10. F. Vincent, C. Bowler, Diatoms are selective segregators in global ocean planktonic communities. *mSystems* **5**, e00444-19 (2020).
11. G. Drebes, S. F. Kühn, A. Gmelch, E. Schnepf, Cryothecomonas aestivalis sp. nov., a colourless nanoflagellate feeding on the marine centric diatom Guinardia delicatula (Cleve) Hasle. *Helgolander Meeresunters.* **50**, 497–515 (1996).
12. E. E. Peacock, R. J. Olson, H. M. Sosik, Parasitic infection of the diatom *Guinardia delicatula*, a recurrent and ecologically important phenomenon on the New England shelf. *Mar. Ecol. Prog. Ser.* **503**, 1–10 (2014).
13. C. Szegedy *et al.*, "Going deeper with convolutions" in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2015), pp. 1–9.
14. D. W. Henrichs, S. Anglès, C. C. Gaonkar, L. Campbell, Application of a convolutional neural network to improve automated early warning of harmful algal blooms. *Environ. Sci. Pollut. Res.* **28**, 28544–28555 (2021).
15. R. Fuchs *et al.*, Automatic recognition of flow cytometric phytoplankton functional groups using convolutional neural networks. *Limnol. Oceanogr.: Methods* **20**, 387–399 (2022).
16. J. J. Suca *et al.*, Sensitivity of sand lance to shifting prey and hydrography indicates forthcoming change to the northeast US shelf forage fish complex. *ICES J. Mar. Sci.* **78**, 1023–1037 (2021).
17. S. M. Lucey, M. J. Fogarty, Operational fisheries in New England: Linking current fishing patterns to proposed ecological production units. *Fish. Res.* **141**, 3–12 (2013).
18. Z. Chen *et al.*, Long-term SST variability on the Northwest Atlantic Continental shelf and slope. *Geophys. Res. Lett.* **47**, e2019GL085455 (2020).
19. J. A. Yoder, S. E. Schollaert, J. E. O'Reilly, Climatological phytoplankton chlorophyll and sea surface temperature patterns in continental shelf and slope waters off the northeast U.S. coast. *Limnol. Oceanogr.* **47**, 672–682 (2002).
20. C. B. Mouw, J. A. Yoder, Primary production calculations in the Mid-Atlantic Bight, including effects of phytoplankton community size structure. *Limnol. Oceanogr.* **50**, 1232–1243 (2005).
21. X. Pan, A. Mannino, H. G. Marshall, K. C. Filippino, M. R. Mulholland, Remote sensing of phytoplankton community composition along the northeast coast of the United States. *Remote Sens. Environ.* **115**, 3731–3747 (2011).
22. Z. Zang *et al.*, Spatially varying phytoplankton seasonality on the Northwest Atlantic shelf: A model-based assessment of patterns, drivers, and implications. *ICES J. Mar. Sci.* **78**, 1920–1934 (2021), 10.1093/icesjms/fsab102 (15 June 2021).
23. B. W. Ibelings *et al.*, Chytrid infections and diatom spring blooms: Paradoxical effects of climate warming on fungal epidemics in lakes. *Freshwater Biol.* **56**, 754–766 (2011).
24. A. C. Thomas *et al.*, Seasonal trends and phenology shifts in sea surface temperature on the North American northeastern continental shelf. *Elementa: Sci. Anthropocene* **5**, 48 (2017).
25. A. S. Gsell, L. N. de S. Domis, E. van Donk, B. W. Ibelings, Temperature alters host genotype-specific susceptibility to chytrid infection. *PLoS One* **8**, e71737 (2013).
26. B. J. Callahan *et al.*, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
27. R. K. Shearman, S. J. Lentz, Long-term sea surface temperature variability along the U.S. east coast. *J. Phys. Oceanogr.* **40**, 1004–1017 (2010).
28. M. T. Burrows *et al.*, The pace of shifting climate in marine and terrestrial ecosystems. *Science* **334**, 652–655 (2011).
29. P. G. Verity, V. Smetacek, Organism life cycles, predation, and the structure of marine pelagic ecosystems. *Mar. Ecol. Prog. Ser.* **130**, 277–293 (1996).
30. J. T. Allen *et al.*, Diatom carbon export enhanced by silicate upwelling in the northeast Atlantic. *Nature* **437**, 728–732 (2005).
31. A. E. S. Kemp, J. Pike, R. B. Pearce, C. B. Lange, The "Fall dump"–a new perspective on the role of a "shade flora" in the annual cycle of diatom production and export flux. *Deep Sea Res. Part II: Top. Stud. Oceanogr.* **47**, 2129–2154 (2000).
32. A. Z. Worden *et al.*, Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science* **347**, 1257594 (2015).
33. J. J. Pierella Karlusich *et al.*, Global distribution patterns of marine nitrogen-fixers by imaging and molecular methods. *Nat. Commun.* **12**, 4160 (2021).
34. M. L. Brosnahan *et al.*, Rapid growth and concerted sexual transitions by a bloom of the harmful dinoflagellate Alexandrium fundyense (Dinophyceae). *Limnol. Oceanogr.* **60**, 2059–2078 (2015).
35. R. J. Olson, A. Shalapyonok, D. J. Kalb, S. W. Graves, H. M. Sosik, Imaging FlowCytobot modified for high throughput by in-line acoustic focusing of sample particles. *Limnol. Oceanogr. Methods* **15**, 867–874 (2017).
36. E. C. Orenstein *et al.*, Semi- and fully supervised quantification techniques to improve population estimates from machine classifiers. *Limnol. Oceanogr.: Methods* **18**, 739–753 (2020).
37. P. González *et al.*, Automatic plankton quantification using deep features. *J. Plankton Res.* **41**, 449–463 (2019).
38. E. C. Orenstein *et al.*, Machine learning techniques to characterize functional traits of plankton from image data. *Limnol. Oceanogr.* **67**, 1647–1669 (2022).
39. R. J. Olson, H. M. Sosik, A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot. *Limnol. Oceanogr.: Methods* **5**, 195–203 (2007).
40. J. Prezioso, Cruise results GU 21-02 spring ecosystem monitoring cruise report (2022). https:/doi.org/10.25923/4W08-1D30 (14 February 2023).
41. L. W. Hugerth *et al.*, Systematic design of 18S rRNA gene primers for determining eukaryotic diversity in microbial consortia. *PLoS One* **9**, e95567 (2014).
42. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
43. L. Guillou *et al.*, The protist ribosomal reference database (PR2): A catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**, D597–D604 (2013).
44. O. Russakovsky *et al.*, ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
45. R. Pawlowicz, M_Map: A Mapping Package for MATLAB (2020). version 1.4m, [Computer software], available online at www.eoas.ubc.ca/~rich/map.html.
46. C. A. Greene *et al.*, The climate data toolbox for MATLAB. *Geochem. Geophys. Geosyst.* **20**, 3774–3781 (2019).
47. D. Catlett, E. E. Peacock, E. T. Crockford, H.M. Sosik, Abundance and parasitoid infection *dynamics* of Guinardia delicatula on the Northeast U.S. Shelf from 2006 to 2022 determined by Imaging FlowCytobot. ver 1. Environmental Data Initiative. https://doi.org/10.6073/pasta/7c43f9e037325e70ee736b7782327dd8. Accessed 26 May 2023.