Dose-dependence and small-scale variability in responses to ocean acidification during squid, Doryteuthis pealeii, development

Dose-dependence and small-scale variability in responses to ocean acidification during squid, *Doryteuthis pealeii,* development

Cite this article as: Casey Zakroff, T. Aran Mooney and Michael L. Berumen, Dose-dependence and small-scale variability in responses to ocean acidification during squid, *Doryteuthis pealeii,* development, Marine Biology https://doi.org/10.1007/s00227-019-3510-8

Dose-dependence and small-scale variability in responses to ocean acidification during squid, *Doryteuthis pealeii,* development

Casey Zakroff[1,2,3], T. Aran Mooney[2], Michael L. Berumen[3]

Corresponding Author: Casey Zakroff, czakroff@whoi.edu, (508) 289-3063

ORCID; CZ: 0000-0001-6979-1857, TAM: 0000-0002-5098-3354, MLB: 0000-0003-2463-2742

[1]Massachusetts Institute of Technology-Woods Hole Oceanographic Institution Joint Program in Oceanography/Applied Ocean Science and Engineering, Cambridge, Massachusetts, USA

[2]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

[3]Red Sea Research Center, Division of Biological and Environmental Science and Engineering, King Abdullah University of Science and Technology, Thuwal, 23955-6900, Saudi Arabia

Keywords: cephalopod, embryo, hypercapnia, paralarvae, statolith, stress

Abstract

Coastal squids lay their eggs on the benthos, leaving them to develop in a dynamic system that is undergoing rapid acidification due to human influence. Prior studies have broadly investigated the impacts of ocean acidification on embryonic squid, but have not addressed the thresholds at which these responses occur or their potential variability. We raised squid, *Doryteuthis pealeii* (captured in Vineyard Sound, Massachusetts, USA: 41° 23.370N 70° 46.418′W), eggs in three trials across the breeding season (May - September, 2013) in a total of six chronic pCO_2 exposures (400, 550, 850, 1300, 1900, and 2200 ppm). Hatchlings were counted and subsampled for mantle length, yolk volume, hatching time, hatching success, and statolith morphology. New methods for analysis of statolith shape, rugosity, and surface degradation were developed and are presented (with code). Responses to acidification (e.g., reduced mantle lengths, delayed hatching, and smaller, more degraded statoliths) were evident at ~ 1300 ppm CO_2. However, patterns of physiological response and energy management, based on comparisons of yolk consumption and growth, varied among trials. Interactions between pCO_2 and hatching day indicated a potential influence of exposure time on responses, while interactions with culture vessel highlighted the substantive natural variability within a clutch of eggs. While this study is consistent with, and expands upon, previous findings of sensitivity of the early life stages to acidification, it also highlights the plasticity and potential for resilience in this population of squid.

Introduction

Addressing the potential effects of ocean acidification (OA) has become a major concern for the management of coastal ecosystems. This includes the northwest Atlantic coastal region where urban development and freshwater influx can exacerbate decreasing pH caused by anthropogenic carbon dioxide (CO_2) (Gledhill et al. 2015). This ecosystem is home to a suite of fisheries species that use nearshore habitats as breeding grounds. Early life stages are expected to be more sensitive to environmental stress than juveniles or adults, so rapidly intensifying impacts such as acidification are of particular concern (Byrne 2011; Haigh et al. 2015).

Loliginid squids, such as the Atlantic longfin squid, *Doryteuthis pealeii*, are common fixtures in many continental shelf ecosystems. These animals are important commercially, with *D. pealeii* supporting a major New England fishery with 18,000 mt landings in 2016 (NOAA Fisheries 2019).They are also a central support structure for the coastal food web, acting as both prey and predator throughout their life history (Jacobson 2005). During reproduction, adults affix their

encapsulated offspring to the nearshore benthos and the young must develop under whatever conditions occur there, potentially resulting in chronic exposure to stressors such as acidification (Jacobson 2005; Fabry et al. 2008; Byrne 2011).

The Atlantic longfin squid comes inshore along the northwest Atlantic coastline from May - October to breed, producing clusters or "mops" of encapsulated embryos that are bound to benthic structure or substrate (Jacobson 2005). Egg laying habitat along the North American Atlantic shelf has been observed to occur at depths shallower than 50 m in salinities of 30-32 ppt and temperatures ranging from 10-23 ºC (McMahon and Summers 1971; Jacobson 2005). Reported shelf carbonate system profiles across *D. pealeii* egg laying habitat suggest a potential exposure range of 8.2 - 7.88 pH_t during breeding season (250 - 600 ppm CO_2; values calculated across depth/temperature extremes using CO_2SYS with data from Wang et al. 2013). Whether pH or others oceanographic parameters, such as oxygenation, determine *D. pealeii* egg laying habitat has not been reported to our knowledge, but observations of the California market squid, *Doryteuthis opalescens*, demonstrate a preference for oxygen levels greater than 160 µmol and pH_t greater than 7.8 (Navarro et al. 2018).

The embryos are packaged inside an egg capsule comprised of mucosal proteins with several hundred siblings, all developing and respiring together (Arnold et al. 1974; Jacobson 2005). Under natural conditions, the inside of these capsules become increasingly anoxic and acidic as development proceeds, reaching pH values as low as 7.34 (Gutowska and Melzner 2009; Long et al. 2016). The only energy source available to these embryos for use in growth, development, and homeostasis is the yolk provided by the mother (Arnold et al. 1974; Steer et al. 2004). While cephalopods are adept at maintaining internal pH balance through the activation of proton secreting transporters within ion-transport epithelia, this process is energetically costly (Hu et al. 2010, 2013). Sensitivity to pH, and the associated homeostatic costs, may vary depending on the cephalopod species and the developmental stage as well (Hu et al. 2010, 2011a).

Previous studies have looked broadly at the potential impacts of acidification on developing loliginid squid embryos. Embryos of *Loligo vulgaris*, removed from the egg capsule and exposed to acidification (pCO_2 ~1650 ppm) and warming (+2ºC), demonstrated delays in development as well as a dramatic decrease (47%) in embryonic survival (Rosa et al. 2014a). *Doryteuthis opalescens* egg capsules cultured under decreased pH (pH 7.57, pCO_2 ~1440 ppm) and hypoxia (80 µM O_2) also showed delays in embryogenesis (Navarro et al. 2016). Further, this study suggested that the combination of these stressors, potentially driven by the hypoxia, resulted in smaller embryonic

statoliths, the aragonitic structures responsible for the squid's sensing of balance, orientation, and sound (Navarro et al. 2016). Kaplan et al. (2013) measured *D. pealeii* paralarvae hatching from eggs reared in high acidification (2200 ppm) and observed both a reduction of statolith size and apparent structural degradation, although the latter was only qualitatively defined. This study also noted a delay in development time and a reduction in paralarval dorsal mantle length as a result of the high acidification dose (Kaplan et al. 2013).

While it is becoming apparent that loliginid squid can be influenced by OA, the additional variables and limited pCO_2 concentrations tested in some of the prior studies make it challenging to assess the scope of pCO_2 impacts. To aid management of this key fisheries species, it is crucial to address whether developmental changes occur gradually with increasing OA or if there is some "tipping point" beyond which effects are significant. Studies addressing early life history are critical because these animals form the foundation for future populations and this phase of development may be particularly vulnerable (Byrne 2011). While documenting fundamental OA effects on these squid is necessary, it is also vital to move beyond basal observations of impacts to address how these animals might cope with this stressor, such as through management of the energy budget, and explore the potential for resiliency within a hatchling cohort.

The experiments performed here were designed to expand upon the work of Kaplan et al. (2013) in order to more thoroughly describe the sensitivity of *D. pealeii* to ocean acidification and understand the mechanisms by which it impacts the early development of this species. We reared *D. pealeii* eggs in a range of pCO_2 treatments in order to examine dose-dependent responses under the hypothesis that between the ambient and 2200 ppm treatments used in the original study lie some physiological threshold for OA. Based on the results from Kaplan et al. (2013), we hypothesized that *D. pealeii* compensated for pH stress by slowing development rate and reducing energy spent on growth, however we did not have a sufficiently robust picture of energy physiology to support this idea. We therefore expanded upon the previous analyses of dorsal mantle length, hatching time, and statolith size and quality (quantifiable metrics were developed), and added measurements of yolk volume (to quantify potential energy consumption effects) and hatching success (to address embryonic survival). We also analyzed data at a high resolution, across multiple hatching days in repeated trials over the breeding season, and describe the natural variability, the potential for resiliency, observed in the squid eggs in response to chronic acidification stress.

Materials and Methods

Squid collection and husbandry

Experiments were performed at the Woods Hole Oceanographic Institution Environmental Systems Laboratory (ESL), Woods Hole, Massachusetts, USA from June through August of 2013. Peak breeding season for *D. pealeii,*in this region, when the squid move into the nearshore of New England, typically falls between May and September (Arnold et al. 1974; Jacobson 2005). Squid were captured in Vineyard Sound by trawls performed by the Marine Biological Laboratory (MBL) in 10-30 meters water depth at the Menemsha Bight of Martha's Vineyard, a locally known breeding ground. Adult squid were hand-selected directly from the trawl ship at the dock. Eighteen medium-sized individuals (20-25 cm dorsal mantle length) that did not appear stressed (those calmly hovering or resting at bottom of the holding tank) or damaged (those without fin tears or skin lesions) were carefully transferred to seawater-filled coolers and driven to the ESL. On top of condition, reproductively active females were selected for based on their bright orange accessory nidamental gland, while males with dense sperm packets visible in the posterior mantle were chosen. Transport occurred as immediately (< 6 hours post-capture), expediently, and gently as possible to minimize stress.

At the ESL, squid were transferred from the coolers into two flow-through cylindrical holding tanks (120 cm diameter, 70 cm depth) fed with water pumped directly from Vineyard Sound to the ESL and continuously bubbled with air. Squid were selected and housed in a 2:1 female to male ratio in order to increase the probability of breeding and egg deposition. Ambient Vineyard Sound seawater was sand-filtered and cooled to 15 ºC (Salinity = 33 psu, pH_{nbs} = 7.96). This temperature falls within the range experienced naturally during the breeding season, but below peak summer temperatures for Vineyard Sound (10.2 - 25.8 ºC from May - October 2013 from NOAA Station BZBM3). Compared to maintaining squid at ambient temperatures, maintaining squid at 15 ºC served to reduce metabolic stress and the occurrence of infighting and cannibalism among the squid, which substantially increased the likelihood of successful egg production. Squid were fed killifish, *Fundulus heteroclitus*, caught in local saltwater ponds once to twice per day, depending on demand. All squid were fed and managed in the ESL until they died after breeding. New adult squid were acquired for each trial.

Female squid laid eggs two to three days after being brought to the ESL. The egg capsules of this species of squid are long, orange-tinged fingers housing 90 - 300 eggs, which are tied together with mucosal proteins into mops that are bound to benthic substrate or structures (Arnold et al. 1974; Maxwell and Hanlon 2000). In the morning, tanks were examined and if egg capsules were

discovered they were immediately hand-transferred into a bucket of seawater from the adult tank and carried into the room with the acidification and culture system. Egg capsules of good quality (thin, oblong, tinted orange, and undamaged) were randomly hand-sorted into the experimental culture cups, two egg capsules per culture cup, to initiate a trial (described below).

Ocean acidification system

Seawater was acidified in a flow-through culture system constructed in a separated room within the ESL. Vineyard Sound seawater pumped into the ESL went through the facility's sand-filters and was then subsequently heated to 20 ºC. This temperature represents the average sea surface temperature for Vineyard Sound over the breeding season (19.5 ºC from May - October 2013 from NOAA Station BZBM3) and resulted in a consistent fourteen day development period for the squid embryos under control conditions. The heated seawater then went through an additional 10 µm filter (Hayward FLV Series, 10 µm felt bag, Hayward Industries, Inc., Rockville, Maryland, USA) to limit small zooplankton, particulate matter, and algae. The water was further treated with a UV sterilizer (Emperor Aquatics Smart HO UV Sterilizer, Model 025150, Pentair Aquatic Eco-Systems, Inc., Cary, North Carolina, USA), in order to reduce harmful protozoa, although flow rate was too high for the seawater to be completely sterilized of microorganisms.

The resultant cleaned and heated water was then output into the header tank of the acidification system, which was vigorously bubbled with compressed air. Between the filtration and heating systems of the ESL and this system header tank, it is expected that most input seawater is mixed over the course of several hours and is not subject to small-scale environmental variability, however fluctuations, particularly of alkalinity, were possible. Fine temporal scale water quality measurements were not performed. Water flowed out of the header into four H-shaped PVC gas equilibration chambers (Fig. S1).. Two air stones in each leg of the 'H' of an equilibration chamber bubbled the flowing seawater with the treatment mixture of compressed air and CO_2. During the first two trials in July (Jul 3 & Jul 11; Table 1) it was discovered that the ambient seawater in the ESL had an elevated concentration of equilibrated CO_2: 550 ppm in the facility compared to 400 ppm for seawater samples taken at depth at the pump intake in Vineyard Sound (carbonate system measurements analyzed with VINDTA). Subsequently, the ambient treatment line of the equilibration chamber section of the acidification system was rebuilt to include two additional chambers, resulting in a line wherein the water was first degassed by N_2 before being re-equilibrated with ambient compressed air in the following two chambers.

Gas mixtures were produced by combining compressed air, introduced at 30 psi from an air compressor within the ESL, with cylinder CO_2. The compressed air was split using a six-way manifold in order to provide aeration through the air stones in the header tank and the equilibration chambers, feed the manifold providing gas to the control culture cups, as well as feed three mass flow controllers (GFC17, Aalborg, Orangeburg, New York, USA), which brought the flow rate down to 4.5 l min^{-1}. Carbon dioxide was also delivered at 30 psi to a parallel set of three mass flow controllers (GFC37, Aalborg, Orangeburg, New York, USA), which were adjusted in order to produce the desired concentrations of CO_2. The air and CO_2 lines were joined downstream of the mass flow controllers and these mixtures were then fed into manifolds which split the gas between the air stones in the equilibration chambers and the bubblers in the culture cups of each treatment. A CO_2 analyzer (model s151, Qubit Systems, Kingston, Ontario, CA), 3-point calibrated with three reference gases (cylinders with 0, 362, and 1036 ppm CO_2, Corp Brothers, Inc., Providence, Rhode Island, USA), was used to check CO_2 concentration in the gas mixtures prior to each trial.

Treatment water flowed from the equilibration chambers into four PVC manifolds from which individual drip lines were connected to the individual culture cups. Egg capsule culture cups were constructed from 1-liter PET food service containers (Solo Foodservice, Lake Forest, IL), which had been pre-soaked in seawater for at least 24 hours and cleaned with deionized water to remove any residues or toxins. These cups had a small rectangular outflow window (2 x 4 cm) cut high on the side and screened with 5 µm mesh, which retained the hatched paralarvae. Each cup was sealed with a lid pierced with two holes, one for the treatment drip line and one for a gas line to bubble in the treatment gas mixture (Fig. S1). Drip lines were fed to the bottom of the culture cup to ensure mixture and overturn and prevent waste accumulation. Treatment water inflow was maintained at a rate of approximately 20 L day^{-1} in each cup, which allowed for sufficient time for the water to equilibrate within the H-shaped chambers. The bubbling line was placed approximately midway under the screened outflow window in order to circulate the water without disturbing the egg capsules while also pushing resulting hatchlings away from the screen. Water from the culture cups outflowed into a communal water bath maintained at 20 ºC using both aquarium chillers (Oceanic Aquarium Chiller 1/10hp, Oceanic Systems, Walnut Creek, California, USA) and a set of controllable aquarium heaters (JÄGER 3603, EHEIM GmbH and Co., Deizisau, DE).

The system consisted of two water baths, allowing for two staggered trials to be run simultaneously (Fig. S1).. Each water bath housed three acidification treatments with four culture cups each for a total of twelve cups. Three cups per treatment were used to culture egg capsules,

while the fourth was used as an abiotic control to monitor water chemistry. An Onset HOBO data logger (HOBO pendant model UA-004-64, Onset Data Loggers, Bourne, Massachusetts, USA) was placed in each water bath to monitor temperature and ambient light; recordings were taken every 15 minutes. Water bath 1 had a mean temperature of 20.49 ± 0.69 °C (mean ± standard deviation) and water bath 2 had a mean temperature of 20.26 ± 0.49 °C across experiments (Table 1). The egg capsules did not undergo temperature acclimation during the transfer from the 15 ºC holding tank to the 20 ºC culture cup, as this level of temperature shift at this early stage of development was not seen to impact embryonic development or survival, or the morphology and physiology of the paralarvae, within the metrics measured here. Ceiling mounted fluorescent lighting in the ESL room containing the culture system was set to a 14:10 light:dark photoperiod to reflect the average natural light cycle for the region. The system was allowed to run for several days prior to acquiring squid in order to ensure equilibration of CO_2 and temperature and balancing of gas bubbling and water flow.

The pH_{nbs} of all of the culture cups, both with and without egg capsules, was monitored by taking samples every three days and measuring using a pH probe (Orion Star™ A329, Thermo Fisher Scientific Inc., Waltham, Massachusetts, USA). These measurements were not considered an accurate proxy for seawater pCO_2, but were used to regularly check pH stability within the system. Respiration of the egg capsules did not notably change pH of the experimental cups compared to the procedural controls. On the day a trial began, and every seven days after (twice more overall), water samples from the fourth cups, the abiotic controls, were taken for high precision carbonate chemistry measurements. pH_t, salinity, and alkalinity data were recorded following the methods adapted from White et al. (2013). In brief, pH_t was recorded using 2 mM *m*-cresol purple indicator dye in a spectrophotometer (USB4000 Spectrometer, Ocean Optics, Dunedin, Florida, USA) using methodology adapted from Clayton and Byrne (1993) and Dickson et al. (2007). Parallel to the pH_t readings, 120 mL glass bottle samples were taken for salinity measurement, which were later analyzed using a Guildline model 8400B "Autosal" laboratory salinometer (Guildline Instruments, Smith Falls, Ontario, Canada). For total alkalinity, 20 mL acid-washed, glass scintillation vials were filled with treatment seawater and poisoned with 10 µL saturated mercuric chloride before being sealed for later analysis. One mL subsamples were processed in duplicate on an automated small volume titrator (Titrando 808, Metrohm AG, Herisau, CH) programmed to run Gran titrations with 0.01 N HCl. ESL seawater samples of known alkalinity, measured using a VINDTA (marianda, Kiel, DE), were used as calibrating standards. If duplicate sets had a difference between samples of 4 µmol kg^{-1} seawater or greater, a second duplicate set was run and the average of the four values

was used. The high precision carbonate chemistry measurements (pH_t, salinity, and total alkalinity), as well as the water bath temperature readings, were input into CO2SYS (Pierrot et al. 2006). Dissociation constants from Mehrbach et al. (1973) and sulfate constants from Dickson (1990) were used in order to calculate pCO_2 and aragonite saturation state (Ω_{arag}) for the equilibrated seawater of each acidification treatment (Table 1).

Experimental trials and paralarval sampling

Trials were initiated by the morning discovery of a mop of egg capsules in the adult holding tanks and are referred to by this laying date throughout this analysis (Table 1). Egg capsules were immediately transferred by seawater bucket(s) to the culture/acidification system room. There they were randomly hand-sorted into the culture cups of an available water bath, two egg capsules per experimental cup (eighteen total egg capsules per trial). Because *D. pealeii* females store sperm and often mate with multiple males, and given that multiple females will lay egg capsules together in the same mop of eggs, the egg capsules used here are of distinctly complex and unknown parentage (Hanlon and Messenger 1998; Buresch et al. 2009). Thus, measurements of adult squid size and weight were not taken. At most, these egg capsules can be considered to represent a haphazard (since size and condition were considered during selection from the trawl catch) sampling of the Vineyard Sound population at a particular point in time during the breeding season. The August trial (Aug 7), however, was initiated by the discovery of egg mops in both of the adult squid holding tanks on the same morning. As a result, egg capsules were randomly distributed, but into a specific set of cups, such that the first cup of each treatment contained two egg capsules from holding tank A, the second cup contained two egg capsules from holding tank B, and the third cup contained one capsule from each tank. Discrete separation of egg capsule parentage therefore occurred for this trial (Cup 1: AA, Cup 2: BB, Cup 3: AB); as much as is possible excepting the probability that a female from one tank had stored the sperm of the male of another tank while in the wild or during capture.

A total of three trials were performed using six carbon dioxide concentrations between 400 ppm and 2200 ppm (Table 1; 400, 550/ESL Ambient, 850, 1300, 1900, and 2200 ppm). The Jul 3 trial was designed to repeat the levels used in Kaplan et al. (2013), atmospheric ambient and 2200 ppm, and add a 1300 ppm midpoint. As stated above, the discovery of the elevated pCO_2 in the ESL seawater affected the ambient treatment of this first trial, and so is reported as ESL Ambient (550) rather than an atmospheric concentration control. Vertical profiles and water column bottle

samples taken in the Menemsha Bight in Vineyard Sound in 2014 and 2015 indicate a consistently well-mixed system with near atmospheric CO_2 concentrations from which this ESL Ambient deviates, but not greatly (July - September 2014, mean bottom [20 m] pCO_2 of 471.8 ppm and average surface to bottom difference of 2.2 ppm; May - September 2015, mean bottom pCO_2 of 484.9 ppm and average surface to bottom difference of 12.7 ppm; Zakroff & Mooney, unpublished data). The Jul 11 trial was intended to be run with ambient control, 850 ppm, the midpoint between 400 and 1300, and 2200 ppm. It was instead run without an ambient control as that treatment line was under reconstruction, and the still active 1300 ppm line was included in its place. Two trials were planned to follow the reconstruction of the 400 ppm line, one using 1600 ppm and one using 1900 ppm, in order to evenly fill the space between 1300 and 2200 ppm. The trial incorporating the 1600 ppm CO_2 treatment was lost, however, due to a failure of the compressed air system resulting in extended exposure of the egg capsules to degassed, and thus deoxygenated, water and is not reported here. The Aug 7 trial was run successfully with the appropriate 400 ppm CO_2 control, representing present atmospheric concentrations, in place, the 2200 ppm treatment acting as the consistent concentration measured across all trials, and the 1900 ppm treatment as planned.

The squid egg capsules were monitored, but otherwise left to develop undisturbed in their culture cups within the acidification system. At eleven to twelve days into development, morning checks for premature hatching began. Hatching typically initiated in the ambient treatment thirteen to fourteen days into development, as expected. Once hatching began, paralarvae were subsampled for a range of experiments. The results of the developmental and morphometric analyses, described below, are reported here, while concurrent behavioral work that subsampled paralarvae from these same experiments has been reported separately (Zakroff et al. 2018).

Dorsal mantle length

Ten paralarvae (fewer if fewer were available) were subsampled from each cup of each treatment each day for the first four to six days of hatching (dependent on hatching dynamics) to be photographed for dorsal mantle length (DML) measurement. When counting out the animals, individuals were pipetted from their culture cup into their own wells within a 24-well plate (Falcon® Brand 2.0 cm^2 well area, 3.5 mL well volume, Corning Inc., Corning, New York, USA). Wells were filled with the appropriate treatment seawater with a few drops of 7.5 % w/v $MgCl_2$ mixed with equal part seawater added as an anesthetic. Individuals were then carefully pipetted into a drop of treatment seawater on a watch glass and placed under a dissecting scope (SteREO

Discovery.V8, Carl Zeiss AG, Oberkochen, DE). Once a paralarva was confirmed to be oriented with the dorsal surface up, most easily recognized on *D. pealeii* by the hexagonal pattern of chromatophores on the dorsal surface of the head, an attached camera (G12, Canon USA, Melville, New York, USA) was used to take a photograph (Fig. 1A). Prior to taking sample photographs, the dissecting scope was focused using the first paralarva and then a calibrating photograph of a millimeter ruler was taken for that day of data collection. If for any reason the focus needed to be changed or photography was interrupted and the camera had to be reset, a new calibration photo was taken. No premature paralarvae, those with remnant external yolk present, nor any that had damage to the mantle, were included in the DML photography dataset. The images of the subsampled paralarvae were later measured for DML using ImageJ (National Institutes of Health, Rockville, Maryland, USA).

Yolk sac volume

An additional 10 paralarvae (fewer if fewer were available) from each cup of each treatment were subsampled each day for lipid staining and preservation using methods adapted from Gallager et al. (1986). In brief, subsampled paralarvae were pooled by treatment cup in a 24-well plate and then euthanized with an increasing dose of 7.5 % w/v $MgCl_2$ mixed with equal part seawater. The paralarvae were then quickly fixed with 10% formalin in order to prevent contraction of the mantle during staining and preservation. The seawater containing the $MgCl_2$ and formalin was then pipetted off and the fixed paralarvae were submersed in oil red O suspended in ethylene glycol, covered, and left to stain overnight. The subsequent morning, the stain was pipetted off and the paralarvae underwent two 30-min soaks in ethylene glycol to remove excess stain before being stored in ethylene glycol in labeled microcentrifuge tubes (0.65 mL Costar microcentrifuge tubes, Corning, Inc., Corning, New York, USA). No notable shrinkage as a result of euthanasia, brief formalin fixation, lipid stain, or ethylene glycol storage was observed, however this was not robustly measured, so analyses are reported under the assumption of either no shrinkage or consistent shrinkage across samples.

Oil red O effectively stained the interior yolk sacs, making them much more visible through the translucent mantle. Lipid-stained paralarvae were photographed in daily sets as described for the DML subset above, except that paralarvae were oriented either dorsal or ventral dependent on which produced the clearer image of the yolk sacs (preferentially ventral, but occasionally dorsal as in Fig. 1B). The photos were processed in ImageJ following the methods of Vidal et al. (2002). In

brief, lines were drawn measuring the length and width of both the anterior and posterior yolk sacs (Fig. 1B). These values were input into formulas representing a three-dimensional shape approximating the volume of the yolk sac: a cone or cylinder for the anterior, and a rotational ellipsoid for the posterior. These results were then summed to get the total yolk volume (YV) for each individual paralarva.

Hatching time and success

Following all morphometric and behavioral work, the remaining paralarvae were counted as they were pipetted from a treatment cup into a petri dish containing water of the same treatment. Paralarvae were then anesthetized with 7.5 % w/v $MgCl_2$ mixed with equal part seawater and preserved in 97% ethanol in microcentrifuge tubes (0.65 mL and 1.7 mL Costar microcentrifuge tubes, Corning, Inc., Corning, New York, USA) by treatment cup (the anesthetized DML subsamples were added back to their appropriate tube in preservation). Thus, no hatched animals remained, nor were any returned to their treatment cups at the end of each experiment day, and all paralarvae analyzed were from their day of hatching (less than 24 hours old). The total number of hatched paralarvae was summed at the end of hatching and used to calculate percent hatching in each treatment cup per day.

Hatching was considered finished in a treatment cup following two mornings with no newly hatched paralarvae. The two egg capsules within the treatment cup would then be removed, photographed, and dissected underneath a dissecting scope. Unhatched embryos were sorted and counted according to their stage of development, adapted from Arnold et al. (1974): early (stages 1 - 16), middle (stages 17 - 26), and late (stages 27 - 30). The total number of unhatched embryos was summed with the total number of hatched paralarvae to get the original count of embryos in each cup. The ratio of hatched paralarvae was compared with the total number of embryos to examine hatching success within each treatment cup.

Statolith morphometrics

Ethanol preserved paralarvae were dissected for their statoliths. An individual paralarva was placed on a glass cover slip underneath a dissecting scope and kept moist with 97% ethanol. Once separated, an individual statolith was rinsed with 97% ethanol and all visible adhering tissue was removed. The statolith was then transferred onto a sticky carbon pad (C249/N 12 mm

diameter self-adhesive carbon disc, TAAB Laboratories Equipment Ltd., Berks, England, UK) on a scanning electron microscopy (SEM) stub (12.7mm diameter aluminum mount, Electron Microscopy Sciences, Hatfield, Pennsylvania, USA). Only one, randomly chosen statolith from each individual paralarva was mounted for SEM imaging and approximately five statoliths (more if possible) per treatment cup (approximately fifteen statoliths per acidification treatment) were assessed. The SEM stubs were taken to the MBL Central Microscopy Facility where they were sputter-coated with 10 nm platinum and imaged using a Zeiss NTS Supra 40VP (Carl Zeiss AG, Oberkochen, DE).

SEM images (1024 x 768 px, TIFF files) were resized such that all statoliths were set to the same 6 px μm^{-1} scale (approximately 672X magnification) using Adobe Photoshop (Adobe Photoshop CC 2017, Adobe Systems Inc., San Jose, California, USA). The Photoshop quick selection tool was then used to select the statolith, carefully maintaining edge integrity. The selection was then cut to a new layer and that layer was saved separately as a PNG for the MATLAB surface analysis described below. The process was then backed up to the selection step and the selection was flood-filled black and again cut to a new layer. Statoliths were then reoriented such that the longest axis of the statolith was parallel to the horizontal axis of the image, the dome, the wider, lobe-like part of the structure, was placed to the right of the image and the rostrum, the thinner, wing-like projection, was to the left (note in Fig. 2B, D that processing in Momocs, described below, flipped these so that the outline had the dome oriented left). For degraded or misshapen statoliths, a best approximation was used, with the longest axis being set horizontal and the subsequently wider side set to the right. The background layer was then flood-filled white to create a black silhouette of a statolith on a white background. These silhouetted statolith images were saved as JPEG files and compiled with the rest of the samples for import into the R (version 3.3.3, R Foundation for Statistical Computing, Vienna, AT) morphometrics package Momocs (Bonhomme et al. 2013) in RStudio (version 1.0.136, RStudio, Inc., Boston, Massachusetts, USA). Momocs took the silhouetted images and translated them into objects describing the outlines of the 2D shapes. Morphometric analysis of the outlines provided statolith length, width, surface area, rectangularity, and circularity.

Two metrics were developed in order to quantitatively describe and compare qualitative observations of statolith degradation. The first was intended to describe the 'rugosity' of the statolith edge, e.g. whether the statolith had a smooth perimeter (Fig. 2A) or a rough one (Fig. 2C). Momocs describes a 2D shape via a series of points that demarcate its outline. Through extensive

testing with a test set of shapes (described in the Supplementary Materials), it was found that calculating the variation of the internal angles between points on an outline at a resolution of 150 points resulted in the best description of shape outline complexity or 'rugosity.' Internal angle variance was calculated for all statolith outlines at this resolution using the Momocs objects in R (code available in the Supplementary Materials and at https://github.com/czakroff/Statoliths).

The second metric was intended to quantify the consistency of the visible statolith surface in the SEM image, e.g. whether a statolith had a smooth surface with organized calcium carbonate crystals (Fig. 2A) or had a rough surface due to increased porosity or disorganized crystals (Fig. 2C). This was achieved by analyzing the average variance of the pixel intensities in five boxes haphazardly placed on the statolith image. The scaled cutout statolith PNG images described above were loaded into a custom MATLAB (version R2016b, Mathworks, Inc., Natick, Massachusetts, USA) script (available in the Supplementary Materials and at https://github.com/czakroff/Statoliths). The script requires the user to click to mark the centroids of five 100 x 100 px squares (equivalent to 277.78 μm^2 of the statolith surface at this scale) that are placed on the statolith image (Fig. 2A, C). The user can then iterate through this process to adjust the squares to ensure they are placed appropriately. Squares were placed in order to achieve as even a distribution over the statolith surface as possible while trying to avoid surface occlusions (salt crystals or remnant tissue), dramatic lighting gradients, and large cracks or breaks. The pixel variance of each box was calculated and then the average surface pixel variance over the five boxes was compiled for all sample statoliths.

Statistics

All statistical analyses were performed in Python (version 3.5.2, Python Software Foundation) using Jupyter Notebooks (Project Jupyter). All data, at all levels (trial, treatment, date, and cup), was tested for normality using Shapiro-Wilk tests ($\alpha = 0.05$) and through visual assessment of quantile plots and histograms. Differences in data that were normally distributed were tested with multi-factor Type II ANOVAs. Type II ANOVAs were selected in order to test factors independently, without ordering, and to not test a main effect in light of its interactions; the hypotheses tested are therefore if a factor in all of its forms has an impact on the dependent variable (Langsrud 2003). Under this framework, the presence of an interaction is of greater interest than a main effect. Treatment (pCO_2) and date were considered independent factors nested under trial, while cup was considered nested under pCO_2. Significant ($P < 0.05$) results were further

analyzed using a Tukey's HSD posthoc test. Statistics of normally distributed data are reported as means ± standard deviation. Yolk sac volume data was log transformed to stabilize variance and then tested as with other parametric data; yolk volume data is reported as the mean and values ± one standard deviation back transformed.

Nonparametric data were assessed for differences using a Kruskal-Wallis test (KW); significant ($P < 0.05$) results were further analyzed using a Dunn's posthoc test. Nonparametric statistics are reported as medians and interquartile range. Distributions of hatching and embryo counts were compared using G-tests. Scatterplots of data by trial are presented with trend lines, primarily as a visual aid. These lines were assessed by linear,regression (LR), significance ($\alpha = 0.05$), but as they were run on three data points we are not suggesting they are statistically powerful. Data compiled across trials was corrected for trial variability (likely due to variability from season/cohort/parentage) by taking the differences between samples and the trial mean, allowing for a comparison of effect size/response slopes (see note on assumptions below). Compiled data were assessed by piecewise linear regression (minimum of three data points per regression) and the model with the best fit (highest R^2) is presented. Piecewise regressions were tested for significant difference against the null hypothesis model of two lines, with different means, each with a slope of zero using a parametric bootstrap.

Results

Water quality

No significant differences in pH_t or calculated pCO_2 were found between cups of the same treatment or across time in the fourth culture cup (KW, $P > 0.05$ for pH_t and pCO2 for all treatments of all trials). Input gas mixtures of carbon dioxide and air were consistent throughout the experiment, resulting in consistent pH_t values (Table 1). Salinity and temperature also remained constant across a trial, but was slightly more variable depending on the cycling routine and sensitivity of the control chiller (Table 1). Temperature also shifted slightly across trials, likely as an effect of changing local environmental temperatures. Calculated aqueous pCO_2 concentrations varied slightly from input gas concentrations, typically being slightly higher, which could be a result of variability in ESL water alkalinity, inconsistency in the flow rate and equilibration rate of treatment water, or a combination of these factors. Equilibrated pCO_2 variability increased with higher input concentrations, with the strongest acidification treatments being the hardest to maintain a consistent equilibration in. This increased variability is likely a result of choosing a flow-

through egg capsule culture system rather than using a closed or recirculating system. Results are analyzed and reported grouped by input gas concentration rather than calculated aqueous pCO_2 concentrations for concision and clarity.

A note on assumptions

There was significant trial-to-trial variability in the response of these developing squid to ocean acidification stress. Data analysis and figures presented examine data both by individual trial and compiled across trials. We therefore sought to clarify assumptions being made in the analysis and compilation of these data. Due to challenges imposed by the facility, the Jul 11 trial did not have an ambient pCO_2 control. Based on the results across metrics, the similarities between the Jul 3 ESL Ambient and the Jul 11 850 ppm data (effect sizes in mantle length and hatching time between these levels and 2200 ppm, in particular), we chose to include these data in compiled graphics and analyses.

Dorsal mantle length compared across lowest pCO_2 treatments of each trial showed no difference between Jul 3 ESL Ambient and Jul 11 850 ppm, but both had significantly larger paralarvae than the Aug 7 400 ppm clutch (ANOVA, $F(2,492) = 9.874$, $P < 0.001$; Tukey, $P < 0.05$; values reported below). Yolk volume, however, showed no difference between the Jul 11 850 ppm and the Aug 7 400 ppm while both were significantly reduced compared to the Jul 3 ESL Ambient (ANOVA, $F(2,471) = 155.3$, $P << 0.001$; Tukey, $P < 0.05$; values reported below). These shifting baselines, a consequence of seasonal, cohort, and/or maternal effects, must be kept in mind when examining the compiled data for a more generalized population response to acidification.

A note on egg number

During manuscript preparation, a reviewer suggested that the relationship between measured factors (primarily DML and YV) and number of eggs in the egg capsules could be examined for potential trade-offs in maternal investment. As two egg capsules were used per cup, we could at best calculate average egg number for each treatment cup ([# hatchlings + # unhatched eggs] / 2). Rather than showing a trade-off, the data suggested a potential increase in both metrics with increasing number of eggs, although the correlation is much stronger for DML (Fig. S4). While a relationship between initial egg size and hatchling DML has been described (Laptikhovsky et al. 2018), as well as negative correlations with egg density (removed eggs in petri dishes; Villanueva et

al. 2011), a positive correlation between number of eggs and DML or YV has not been reported for multi-egg per egg capsule squids to our knowledge.

Our data represent three egg clutches laid by unknown parents (preselected for size/condition) taken in one breeding season and is not robust enough to consider reevaluating the entire dataset by (particularly without literature support). Type II ANOVAs of DML and YV run with only egg number as an independent covariate and cup (numbered individually rather than nested in treatment) demonstrated no effect of cup ($P >> 0.05$) and a substantial effect of egg number ($P << 0.001$) and a substantial interaction between egg number and cup ($P << 0.001$) in both metrics across trials (Table S5) suggesting that within the scope of our statistical models, these factors represent the same effect (if egg number were categorical, it and cup would be indistinguishable statistically). Statistical models incorporating egg number as an independent continuous covariate are reported in the Supplementary Materials (Tables S7 & S8), but most statistics and data presented here are done so under the assumption that random selection of egg capsules accounted for this potential source of variability.

Dorsal mantle length

Dorsal mantle length decreased with increasing pCO_2 in all trials (Fig. 3). Overall compiled data showed a significant effect of trial and cup on DML, with near significant effects of pCO_2 and the interaction between pCO_2 and hatching date (Table 2). Paralarvae in the Jul 3 trial showed a broadly linear, but non-significant, decrease of DML with increasing pCO_2 (LR, $P = 0.106$). Each treatment in this trial was significantly different from the others (Table S1; Tukey, $P < 0.05$: ESL Ambient/550: 1.64 ± 0.11 mm; 1300 ppm: 1.59 ± 0.12 mm; 2200 ppm: 1.56 ± 0.12 mm). The Jul 11 trial also showed a significant decrease in DML with increasing pCO_2 (Table S1), but showed a non-linear (LR, $P = 0.206$) step-wise response, with the 850 ppm (1.63 ± 0.12 mm) and 1300 ppm (1.63 ± 0.11 mm) treatments being grouped (Tukey, $P < 0.05$) separately from the 2200 ppm treatment (1.53 ± 0.12 mm). Paralarvae in the Aug 7 clutch demonstrated a weaker, but significant reduction in DML with increased acidification (Table S1). Again, a step-wise (LR, $P = 0.123$) response was seen, with the 400 ppm treatment (1.59 ± 0.13 mm) having significantly (Tukey, $P < 0.05$) larger paralarvae than both the 1900 ppm (1.53 ± 0.13 mm) and 2200 ppm (1.53 ± 0.12 mm) treatments. Compiled by difference from trial mean, piecewise regression indicated a two line model, of a low pCO_2/greater DML group and a higher pCO_2/smaller DML group with breakpoint at 1300 ppm, best fit ($R^2 = 0.858$) the data and was not significantly different from the stepwise null hypothesis ($P =$

0.363, Fig. 3). The compiled dataset of differences showed a significant relative decrease in DML with increased acidification (ANOVA, $F_{8,1440} = 16.50$, $P < 0.001$), with statistical groupings splitting at the 1300 ppm treatment (Tukey, $P < 0.05$; Fig. 3).

The significant interaction between pCO_2, hatching date, and cup on DML in the Jul 3 trial had the greatest effect size (Ω^2; Table 2). Interactions between pCO_2 and hatching date alone were also significant (Table 2). The ESL Ambient/550 ppm treatment showed no difference in mean DML across the hatching days (ANOVA $P > 0.05$; Table S2), despite a significantly increasing trend (LR, $P = 0.023$; Fig. 4). The 1300 ppm samples were more variable, with significant differences in DML over the days of hatching (Table S2), but no corresponding trend (LR, $P = 0.780$; Fig. 4). The 2200 ppm exposure approached significance for both mean DML over time and a slight decreasing trend (Table S2; LR, $P = 0.090$; Fig. 4). The effect of cup (nested in treatment) was significant, with a similar effect size to the interaction of pCO_2 and date (Table 2). Details of responses by cup have been placed in the Supplemental Materials for manuscript brevity, except for the Aug 7 trial.

The Jul 11 trial also showed a significant interaction of pCO_2, hatching date, and cup on DML (Table 2). While all factors and interactions showed significant effects on DML, date and the interaction between pCO_2 and hatching date had the greatest effect sizes (Table 2). All three pCO_2 treatments showed significant effects of hatching date on DML (Table S2). While all treatments showed decreasing paralarvae size with time, only the 850 ppm treatment fit a linear trend (LR, $P = 0.007$; Fig. 4). Cup alone was significant, but with a lower effect size than date and its interactions (Table 2).

Differences in DML were much smaller in the Aug 7 trial, but still showing a significant interaction between pCO_2, date, and cup (Table 2;). As with the Jul 11 trial, all factors were significant here, but cup and date had the greatest effect sizes (Table 2). The effect of hatching date on DML was near significance in the 400 ppm treatment and significant in both the 1900 and 2200 ppm treatments (Table S2). All of the pCO_2 treatments demonstrated a non-linear (LR, 400: $P = 0.8632$; 1900: $P = 0.8733$; 2200: $P = 0.5168$) bimodal distribution of DML over hatching, with peaks on the first and fourth days of hatching (Fig. 4).

The Aug 7 trial consisted of egg capsules from two separate adult holding tanks (tank A and tank B) sorted into the culture cups for each pCO_2 treatment (Cup 1: AA, Cup 2: BB, Cup 3: AB). Cup (nested in pCO_2 treatment) had the greatest effect size on DML in this trial, while its interaction had the lowest (Table 2). At the scale of discrimination by cup, egg number could be notably relevant, so values are reported here while detailed statistical analyses can be found in the Supplementary

Materials. In brief, egg number appears to be a significant covariate interacting with all other factors (cup [not nested when acting as a proxy for tank/egg capsule source], pCO_2, and their interaction in particular, Table S7). Integrated across treatments, cup/source had a significant effect on egg number (ANOVA, $F_{2,490} = 284.7$, $P << 0.001$) with all cups/sources being significantly different from each other (Tukey, $P < 0.05$; Cup 1/AA: 128.7 ± 12.1 eggs/capsule; Cup 2/BB: 169.8 ± 19.9 eggs/capsule; Cup 3/AB: 147.0 ± 14.9 eggs/capsule)

Within the 400 ppm treatment, cup had a significant effect on paralarval DML (Table S3) with Cup 2/BB paralarvae significantly (Tukey, $P < 0.05$) larger (1.64 ± 0.12 mm, 192.5 eggs/capsule) than those from Cup 1/AA (1.56 ± 0.12 mm, 117.5 eggs/capsule) and Cup 3/AB (1.54 ± 0.11 mm, 126 eggs/capsule). The 1900 ppm treatment also showed significant differences between cups (Table S3; Tukey, $P < 0.05$), but with Cup 1/AA paralarvae (1.46 ± 0.12 mm, 123 eggs/capsule) much smaller than those from both Cup 2/BB (1.54 ± 0.11 mm, 144 eggs/capsule) and Cup 3/AB (1.59 ± 0.13 mm, 158.5 eggs/capsule). No difference was seen in the 2200 ppm treatment (Table S3; Cup 1/AA: 1.52 ± 0.10 mm, 145.5 eggs/capsule; Cup 2/BB: 1.56 ± 0.12 mm, 173 eggs/capsule; Cup 3/AB: 1.52 ± 0.14 mm, 156.5 eggs/capsule). Paralarvae from Cup 1/AA and Cup 2/BB showed similar patterns of response to the acidification exposure, a non-linear decrease with increased exposure (LR, Cup 1/AA: $P = 0.4789$; Cup 2/BB: $P = 0.2190$), while no trend (LR, $P = 0.9514$) or clear pattern of response was seen in Cup 3/AB (Fig 5). Compiled across cups, the data shows the relative decrease in DML, approaching significance, reported in data by trial (LR, $P = 0.0824$; Fig 3, Fig 5).

Variance of the DML data, assessed by pooling each cup and comparing across treatments within a trial, consistently increased with increasing acidification. No individual t-tests showed significant differences in variance between treatments, likely influenced by low sample size (n = 3, two-sample t(2), $P > 0.05$ for all treatment pairings within each trial). All three trials demonstrated non-significant increasing linear trends in DML variance (Fig. 3; LR, Jul 3: $P = 0.1038$; Jul 11: $P = 0.2297$; Aug 7: $P = 0.1738$). The change in variance, relative to trial average, for all pCO_2 treatments best fit ($R^2 = 0.780$) a two-line model breaking after 1300 ppm (no significant difference from stepwise model, $P = 0.544$; Fig. 3). In the Aug 7 trial, DML variance was highest in Cup 3/AB (0.0159 ± 0.0035 mm^2), but not significantly different from the other cups (two-sample t(2), $P > 0.05$ for all pairings; Cup 1/AA: 0.0129 ± 0.0020 mm^2; Cup 2/BB: 0.0143 ± 0.0011 mm^2).

Yolk sac volume

Patterns of response in yolk sac volume were highly variable within and between trials (Fig. 3). Yolk sac volume in the low/control treatments decreased markedly (one-way ANOVA, $F(2,471) = 166.8$, $P << 0.001$) between the Jul 3 trial (0.077 mm^3, 0.042 - 0.138 mm^3) and the Jul 11 (0.030 mm^3, 0.017 - 0.054 mm^3) and Aug 7 (0.029 mm^3, 0.020 - 0.044 mm^3) trials (Tukey, $P < 0.05$). Despite this, only cup, nested within pCO_2 nested within trial, appears significant when data is compiled (Table 2). In the Jul 3 trial, YV decreased linearly across pCO_2 treatments (LR, $P = 0.017$; Fig. 3) with the ESL Ambient/550 treatment having significantly larger YV (Table S1; Tukey, $P < 0.05$; 0.077 mm^3, 0.070 - 0.084 mm^3) than the 2200 ppm treatment (0.058 mm^3, 0.034 - 0.100 mm^3). Conversely, yolk volume increased near-linearly (LR, $P = 0.060$; Fig. 3) with increasing acidification in the Jul 11 trial with the 850 ppm treatment showing significantly smaller YV (Table S1; Tukey, $P < 0.05$; 0.030 mm^3, 0.017 - 0.054 mm^3) than the 2200 ppm (0.036 mm^3, 0.018 - 0.72 mm^3) treatment. Yolk sac volume was not affected by pCO_2 in the Aug 7 trial (Table S1; Fig. 3). In the compiled data, normalized to trial mean, piecewise regression showed a weakly fitting ($R^2 = 0.221$) two line model, not significantly different from a stepwise null model ($P = 0.839$) with a breakpoint between 850 and 1300 ppm (Fig. 3). Variance of the YV also showed no trends with increasing acidification (LR, $P > 0.05$ for all trials; n = 3, two-sample t(2), $P > 0.05$ for all treatment pairings within each trial) with piecewise regression revealing a two line best fit ($R^2 = 0.609$) model breaking at the lowest values at 1300 ppm (no difference from stepwise null model, $P = 0.304$; Fig. 3).

The interaction of pCO_2, date, and cup had a significant impact on YV in all three trials (Table 2). In the Jul 3 trial, the interaction of pCO_2 with date had the greatest effect size (Table 2) showing trends in YV over hatching similar to the DML data, with the ESL Ambient/550 ppm (LR, $P = 0.140$) increasing slightly, while the 1300 ppm (LR, $P = 0.038$) and 2200 ppm (LR, $P = 0.145$) paralarvae decrease (Fig. 4). All factors and interactions were significant in the Jul 11 trial (Table 2), though weaker than the Jul 3 trial with cup appearing to be a stronger interacting factor with pCO_2 than date. While YV significantly changed with date under the 850 and 2200 ppm treatments (Table S2), no particularly strong trends were seen (LR, $P > 0.05$; Fig. 4). The Aug 7 trial showed no overall effect of pCO_2, but a weak effect of date and near significant interaction of pCO_2 and date (Table 2, Table S1). There were significant differences in YV with date in the 400 and 2200 ppm treatments (Table S2), with all three treatments showing weakly increasing trends over hatching (LR, 400: $P = 0.079$; 1900: $P = 0.069$; 2200: $P = 0.144$; Fig. 4).

Cup and its interaction showed significant effects in the Aug 7 trial (Table 2). Since the correlation of egg number to YV is non-significant and very weak (Fig. S4), statistical models are included in the Supplemental Materials, but are not reported here (Tables S7 & S8). Cup 1/AA showed no significant difference with pCO_2 (ANOVA, $F_{2,153} = 1.32$, $P = 0.268$), but an increasing trend (LR, $P = 0.031$; Fig. 5). Cup 2/BB, conversely showed a significant decrease in YV at the 1900 ppm level (ANOVA, $F_{2,170} = 8.36$, $P < 0.001$; Tukey, $P < 0.05$) driving a slight, but non-significant, decreasing trend (LR, $P > 0.05$; Fig. 5). Cup 3/AB showed no significant effect (ANOVA, $F_{2,148} = 0.315$, $P = 0.730$) and a very weakly increasing trend (LR, $P > 0.05$) in YV with increasing pCO_2 (Fig. 5). The 400 ppm treatment varied by cup (Table S3; Tukey, $P < 0.05$) with Cup 2/BB having greater YV (0.034 mm^3, 0.023 - 0.050 mm^3; 192.5 eggs/capsule) than both Cup 1/AA (0.028 mm^3, 0.019 - 0.040 mm^3; 117.5 eggs/capsule) and Cup 3/AB (0.026 mm^3, 0.017 - 0.038 mm^3; 126 eggs/capsule). Yolk volume variance did not differ between cups of the Aug 7 trial (n = 3, two-sample t test, $t(2)$, $P > 0.05$ for all cup pairings). Based on a comparison of average values for each cup, yolk sac volume was independent of dorsal mantle length (LR, $P > 0.05$, for all trials; Fig. 6).

Hatching time

Increasing acidification consistently delayed hatching in all trials (Fig. 7). Days until hatching initiation, defined as the day at which 1% cumulative hatching occurred in at least one treatment, also increased across trials (Jul 3: 12 days from laying, Jul 11: 14 days, Aug 7: 15 days). In the Jul 3 trial, the proportions of cumulative hatching over time were significantly different between pCO_2 treatments (G-test, ESL Ambient x 1300 ppm: $G(12) = 156.2556$, $P << 0.0001$; ESL Ambient x 2200 ppm: $G(12) = 412.4811$, $P << 0.0001$; 1300 x 2200 ppm: $G(12) = 517.2413$, $P << 0.0001$). Cumulative hatching proportions also varied significantly between cups (G tests, $P < 0.0001$ for all cup pairs within each pCO_2 treatment, except 2200 ppm Cups 1 and 2, $P = 0.2629$). Distributions of cumulative fraction hatched over time, compiled by pCO_2 treatment, are considered here for concision (Fig. 7). Hatching in the 2200 ppm treatment was consistently delayed from the ESL Ambient and 1300 ppm treatments by about 1 day (Fig. 7). Cumulative hatching reached at least 25% at 13 days from laying in the ESL Ambient and the 1300 ppm treatment, but took 14 days in the 2200 ppm treatment. Hatching of 75% or greater was reached 14 days from laying in the 1300 ppm treatment, 15 days in the ESL Ambient treatment, and 16 days in the 2200 ppm treatment.

Proportions of cumulative hatching over time, compiled by pCO_2 treatment, were also significantly different within the Jul 11 trial (850 x 1300 ppm: $G(11) = 81.9224$, $P << 0.0001$; 850 x 2200 ppm: $G(12) = 664.3269$, $P << 0.0001$; 1300 ppm x 2200 ppm: $G(12) = 500.7742$, $P << 0.0001$). Again, some variability in hatching dynamics was seen between cups (G tests, $P < 0.01$ for all cup pairs within each pCO_2 treatment, except 2200 ppm Cups 1 and 2, $P = 0.2880$). Compiled, the distributions show a consistent delay of about 1 day, expanding to 2 days, in the 2200 ppm treatment (Fig. 7). Hatching reached at least 25% 15 days after laying in the 850 and 1300 ppm treatments and 16 days after in the 2200 ppm treatment. Cumulative hatching of at least 75% was reached 16 days after laying in the 850 and 1300 ppm treatments and 18 days after laying in the 2200 ppm treatment.

The Aug 7 trial also showed notable differences between pCO_2 treatments in cumulative hatching proportions over time (400 x 1900 ppm: $G(15) = 693.0624$, $P << 0.0001$; 400 x 2200 ppm: $G(15) = 892.6867$, $P << 0.0001$; 1900 ppm x 2200 ppm: $G(12) = 79.242$, $P << 0.0001$) and between the cups of each treatment (G tests, $P < 0.05$ for all cup pairs within each pCO_2 treatment, except 2200 ppm Cups 2 and 3, $P = 0.0564$). A consistent delay of about 1 day was seen in the 1900 and 2200 ppm treatments, compared to the 400 ppm treatments (Fig. 7). At least 25% cumulative hatching was seen in the 400 ppm treatment 16 days after laying, but 17 days after in the 1900 and 2200 ppm treatments. At least 75% hatching occurred after 18 days in the 400 ppm treatment, and after 19 days in the 1900 and 2200 ppm treatments.

Hatching success

Hatching success was high across pCO_2 treatments and trials, with at least 85% hatching always seen (compiled by treatment; Table S4). No trends in hatching success with increasing acidification were seen in any trial (Jul 3: LR, $P = 0.8199$; Jul 11: LR, $P = 0.2455$; LR, $P = 0.8431$). Significant differences were seen in the distributions of staged, unhatched embryos and hatched paralarvae within treatments and cups in all trials (Table S4; G tests, $P < 0.05$), but followed no pattern with acidification. In the Aug 7 trial, Cup 1 / AA had significantly higher embryonic mortality, particularly of middle and late stage embryos, than both Cup 2 / BB and Cup 3 / AB in both the 400 and 1900 ppm treatments (G tests, $P < 0.05$); no differences were seen in the 2200 ppm treatment (G test, $P > 0.05$ for all cup pairs). Occasional spikes in mortality of early stage embryos (e.g. 30.9% of eggs in Cup 1 of the ESL Ambient / 550 treatment in the Jul 3 trial), either due to natural variability or faults of the culture system, may also have skewed results.

686

687 *Statolith morphometrics*

688 Statolith area broadly decreased with increasing acidification, although responses varied
689 from trial to trial. In the Jul 3 trial, statoliths from the ESL Ambient (6823.8 μm^2, 6449.9 - 7440.3
690 μm^2) treatment were significantly larger (KW, $H2 = 9.0613$, $P = 0.0108$; Dunn, $P < 0.05$) than those
691 from the 1300 ppm (5723.2 μm^2, 5134.2 - 6620.3 μm^2) and 2200 ppm treatments (5803.2 μm^2,
692 5114.0 - 7142.8 μm^2), following an apparent step-wise drop (LR, $P = 0.4090$; Fig. 8). In the Jul 11
693 trial, statoliths from both the 850 (7882.4 μm^2, 7436.3 - 8115.9 μm^2) and 1300 ppm (7778.3 μm^2,
694 7553.5 - 8017.7 μm^2) treatments were much larger (KW, $H2 = 13.9475$, $P = 0.0009$; Dunn, $P < 0.05$)
695 than those in the 2200 ppm treatment (4845.0 μm^2, 3291.6 - 6747.8 μm^2), again following a step-
696 wise drop (Fig. 8; LR, $P = 0.3462$). There was no difference (KW, $H2 = 1.8239$, $P = 0.4017$) in
697 statolith area between pCO_2 treatments in the Aug 7 trial (400: 6814.4 μm^2, 6218.0 - 7074.4 μm^2;
698 1900: 6618.2 μm^2, 5920.2 - 7355.9 μm^2; 2200: 6473.3 μm^2, 6119.3 - 6836.1 μm^2) and no trend in
699 these data (Fig. 8; LR, $P = 0.7102$). Overall, relative to trial, statolith surface area best fit ($R^2 = 0.638$)
700 a two-line model with a breakpoint at 1300 ppm (at which area decreases), which did not
701 significantly differ from a stepwise null hypothesis ($P = 0.681$, Fig. 8). Statolith area appeared to be
702 dependent on mantle length, based on a comparison of average values for each treatment, which
703 approached significance (LR, $P = 0.0519$, Fig. S2).

704 The variance of the internal angle of the statolith outline, the metric of statolith edge
705 rugosity, broadly increased with increasing acidification in the compiled data, driven by the Jul 11
706 samples. Internal angle variance was significantly higher (KW, $H2 = 17.6603$, $P = 0.0001$; Dunn, $P <$
707 0.05) in the 1300 ppm (507.11 deg^2, 412.71 - 715.98 deg^2) treatment of the Jul 3 trial than either
708 the ESL Ambient (225.96 deg^2, 169.39 - 294.59 deg^2) and the 2200 ppm (277.83 deg^2, 130.13 -
709 577.08 deg^2) treatments resulting in a nonsignificant increasing trend (LR, $P = 0.8082$; Fig. 8). In the
710 Jul 11 trial, treatments were not significantly different from each other (KW, $H2 = 4.8128$, $P =$
711 0.0901), but internal angle variance of the statoliths followed an increasing trend with acidification
712 (850: 89.13 deg^2, 63.38 - 367.81 deg^2; 1300: 151.25 deg^2, 88.30 - 308.43 deg^2; 2200: 348.32 deg^2,
713 158.85 - 521.12 deg^2; LR, $P = 0.0252$; Fig. 8). Statolith internal angle variance was much lower
714 overall in the Aug 7 trial, and showed no differences between treatments (KW, $H2 = 4.0206$, $P =$
715 0.1339) and no particular trend with acidification (400: 97.51 deg^2, 79.25 - 115.35 deg^2; 1900:
716 110.49 deg^2, 92.63 - 129.89 deg^2; 2200: 97.44 deg^2, 79.65 - 118.81 deg^2; LR, $P = 0.5197$; Fig. 8). The
717 data compiled relative to trial means best fit a two-line model ($R^2 = 0.716$) with a breakpoint at

1300 ppm (at which internal angle variance increases) that did not differ from a stepwise model (P = 0.277; Fig. 8).

The average variance of statolith surface pixel intensity (px int^2) followed similar patterns as internal angle variance, with a stepwise model (R^2 = 0.573, P = 0.521) increasing at a 1300 ppm breakpoint in the compiled data; again driven by the Jul 11 samples. In the Jul 3 trial, average surface pixel variance was highest in the 1300 ppm (1085.18 px int^2, 854.73 - 1386.61 px int^2) treatment of the Jul 3 trial, significantly above (KW, $H2$ = 13.2045, P = 0.0014; Dunn, P < 0.05) the ESL Ambient (665.77 px int^2, 526.24 - 929.30 px int^2) and 2200 ppm (713.83 px int^2, 448.52 - 849.16 px int^2) treatments (LR, P = 0.8843; Fig. 8). The Jul 11 trial followed a step-wise jump in surface variation (LR, P = 0.2292; Fig. 8), with the statoliths of the 850 (185.41 px int^2, 116.57 - 290.16 px int^2) and 1300 ppm (200.95 px int^2, 155.71 - 282.51 px int^2) treatments having significantly lower surface variation (KW, $H2$ = 16.0099, P = 0.0003; Dunn, P < 0.05) than the statoliths from the 2200 ppm (601.08 px int^2, 422.72 - 691.84 px int^2) treatment. Statolith surface pixel variance was lower in the Aug 7 trial, although it still showed significant differences (KW, $H2$ = 7.9688, P = 0.0186; Dunn, P < 0.05) between pCO_2 treatments, with the 400 ppm treatment (111.24 px int^2, 64.36 - 147.25 px int^2) having lower surface variation than the 1900 ppm (144.26 px int^2, 123.33 - 169.51 px int^2 and 2200 ppm (130.42 px int^2, 100.26 - 178.62 px int^2) treatments, resulting in a weakly increasing trend with acidification (LR, P = 0.1406; Fig. 8).

Rectangularity and circularity of the statoliths were inversely related, demonstrating weak, non-significant trends with increasing acidification (LR, P > 0.05; Fig. 8). Compiled, rectangularity fit (R^2 = 0.426, P = 0.791) a stepwise model decreasing at 1300 ppm. Circularity also best fit a stepwise model (R^2 = 0.657, P = 0.319), but with a breakpoint increasing circularity between 850 and 1300 ppm. In the Jul 3 and Jul 11 trials, where statoliths showed impacts of acidification in other metrics, statoliths appear to become less rectangular and more circular (Fig. 8). Statoliths from the 1300 ppm treatment of the Jul 3 trial had significantly lower rectangularity than those from the ESL Ambient/550 treatment (KW, $H2$ = 17.6603, P = 0.0001; Dunn, P < 0.05; Fig. 8), but this was the only result to support these potential trends, likely a factor of low sample sizes and high variability.

Discussion

This work expands our knowledge of the physiological impacts of ocean acidification on the early development of squid paralarvae, while also demonstrating the capacity for adaptation and

resilience inherent to this fecund, plastic organism. In response to elevated pCO_2, hatchling *D. pealeii* paralarvae demonstrated reduced mantle length, delayed hatching time, and degraded statoliths, consistent with the observations by Kaplan et al. (2013). Breakpoints in the compiled data were consistently around 1300 ppm CO_2 across metrics, although there was notable variability in response strength from trial to trial. This value falls above IPCC predictions for ocean acidification in the open ocean by 2100 (~850 ppm), but below that for 2250 (~1500 ppm), and already occurs naturally, on short time scales, within estuarine and coastal systems (Caldeira and Wickett 2003; Doney et al. 2009; Baumann et al. 2014; Gledhill et al. 2015). Although juvenile and adult *D. pealeii* are known to enter estuarine systems, and thus tolerate some substantive pH variability, the eggs are typically laid in a more stable system: the nearshore shelf bottom up to 50 m depth (Gray 1992; Jacobson 2005). Even at the extremes of observed egg laying habitat, pH_t should not be below 7.8 (about 700 ppm in our system), but developing embryos appear capable of resisting acidification well beyond that mark (McMahon and Summers 1971; Jacobson 2005; Wang et al. 2013). It is likely, as has been observed for *D. opalescens*, that oxygenation delimits egg laying habitat as well as pH (Navarro et al. 2018). Oxygen should not be as restrictive on the Northwest Atlantic shelf, but perhaps for *D. pealeii* oxygen, or still other factors, is a more limiting determinant of the egg laying habitat window than pH. Whereas thermal and hypoxia thresholds are often considered in physiological work, acidification thresholds have primarily been considered for calcifying marine organisms (Anthony et al. 2008; Byrne 2011; Gazeau et al. 2013; Rosa et al. 2013). However, a greater understanding of acidification tolerance windows in more marine organisms could be extremely useful for informing models and producing more robust predictions for fisheries management (Hofmann and Todgham 2010).

Depending on the mechanism through which pH balance is achieved, an organism may reach its limit through either increasing energetic costs or through the accumulation of bicarbonate (Fabry et al. 2008). Cephalopods are highly effective at pH balancing through ion transport, but this process is considered energetically costly (Hu et al. 2011b, 2013). The results presented here indicate an OA threshold for the case of embryonic *D. pealeii*, which have a finite energy reserve, but this ”threshold” may not apply to post-hatch paralarvae and later stages of development which are potentially capable of moving out of stressful pH environments and may supplement energy through feeding (Vidal and Haimovici 1998; Bartol et al. 2008). Similarly, although hatching success was consistently high across trials and treatments, this only acts as a measure of embryonic survival and we cannot make any claims regarding the viability or survival of the resultant hatchlings.

783

784 *Energy budgets under stress: mantle length and yolk reserves*

785 The squid in each trial of this experiment demonstrated a different strategy of energy
786 budget management in response to OA stress. In all cases, development rate was slowed, consistent
787 with the observations of other loliginid embryos under acidification (Kaplan et al. 2013; Rosa et al.
788 2014a; Navarro et al. 2016). It is uncertain if this developmental delay is a result of metabolic
789 depression, which is a common response of marine invertebrates (Pörtner et al. 1998; Michaelidis
790 et al. 2005). While metabolic depression under increased pCO_2 (around 1000 ppm) has been
791 observed in adult Humboldt squid, *Dosidicus gigas*, more recent research indicated that neither
792 adults and juveniles of these squid nor of *D. pealeii* demonstrate metabolic depression or oxygen
793 limitation under hypercapnia (1410 ppm; Rosa and Seibel 2008; Birk et al. 2018). Energy may have
794 been sacrificed from growth in our experiments, as dorsal mantle length decreased with increasing
795 OA in all trials. Yolk volume, however, responded in numerous ways, perhaps a result of varying
796 resiliency, varying coping strategies, or yolk usage being inconsistently affected by pCO_2 level (Fig.
797 6).

798 Comparisons of mantle length and yolk volume highlight potential differences in the
799 response to OA stress across the breeding season. In the Jul 3 trial, both DML and YV decrease
800 slightly with increasing acidification suggesting a stressed system that requires more energy to
801 maintain (Fig. 6). In the Jul 11 trial, DML decreases, but YV slightly increases, as acidification
802 increases suggesting a system of depressed metabolism/growth (Fig. 6). Responses were low in the
803 Aug 7 trial, with YV staying constant as DML slightly decreased with increasing acidification,
804 suggesting either a potentially resilient system or a reduced impact magnitude due to the overall
805 smaller paralarvae in this clutch (Fig. 6).

806 While DML effect size was small, in context of the typical *D. pealeii* paralarvae it accounted
807 for an approximately 5% reduction in size across trials as a result of acidification (integrated over
808 hatching days). Raising *D. pealeii* paralarvae in captivity is a possible, but systemically challenging
809 proposition, so while we unfortunately do not have direct observations of survival in this study we
810 can hypothesize about the multiple pathways through which a reduction of this magnitude could
811 impact the viability and survival of the hatchlings (Vidal et al. 2002b; Steer et al. 2003). The post-
812 hatch transition from consumption of yolk reserves to prey capture is considered a critical period
813 for squid paralarvae, and hatchling size is considered an important factor in prey capture success
814 (Vidal et al. 2002a, b). Further, paralarval hydrodynamics and swimming speeds could be impacted

by shifts in overall size, potentially impairing an already low (40%) ability to escape predation (Bartol et al. 2008; York and Bartol 2016). Yolk volume reduction was seen only in the Jul 3 trial, but showed an average 24% decrease, compounding concerns for paralarval survival of the critical period under that response to acidification stress. Yolk content is also connected to paralarval specific gravity, and has been noted as of potential importance in paralarval survival as part of dispersal (Martins et al. 2010a).

Dorsal mantle length and yolk volume were often strongly affected by hatching date, indicating either natural variability in hatching dynamics and/or an impact of increased exposure time. The latter could be compounded by the delay in hatching time caused by increased acidification. Assuming growth rate for all embryos is consistent and occurs under the same conditions, mantle length would be expected to increase with hatching date, as the embryos that were not triggered to hatch continue to grow (Fig. 4: ESL Ambient / 550). This model of development has been shown in the eggs of bigfin reef squid, *Sepioteuthis lessioniana* (Ikeda et al. 1999). Conversely, seeing a decrease in mantle length as hatching continues indicates embryos that either felt a greater impact of the stressor, lagged in development, and/or lacked in resources (Fig. 4: 2200 ppm; Fig. 4).

We expect that yolk would be consumed as hatching day increased, perhaps to a greater extent for paralarvae under stress. The Aug 7 trial however, broadly showed increases in yolk volume with hatching date in all treatments. Yolk utilization in squid paralarvae is known to be impacted by temperature, driving metabolism, and feeding state (Vidal et al. 2002a; Martins et al. 2010b). Both these factors were consistent across trials and so do not account for the different patterns in yolk utilization seen. Further, assessments of either varying development or yolk utilization rate are confounded by potential differences in maternal ration. Unfortunately, it is not feasible to quantify yolk rations within a capsule without disturbing the embryos and potentially inducing premature hatching. Because maternity was unknown, and potentially mixed, within the egg mops used in each trial, it is possible that differences in maternal investment account for these variable patterns of response across the breeding season (Steer et al. 2004).

Construction of the statolith

Responses of the statolith to acidification followed similar patterns from trial to trial and were fairly consistent across the metrics observed. Statolith length has been correlated to mantle length in squid, so the decrease in statolith area seen with increasing acidification in our data is

likely driven by the concurrent decrease in dorsal mantle length (Fig. S1) (Ikeda et al. 1999; Steer et al. 2003). However, decreases in statolith area due to combined acidification and hypoxia described in *D. opalescens* were independent of paralarval size, so in certain stressor scenarios, statolith and organism size may be decoupled (Navarro et al. 2016).

The increases in statolith edge rugosity and surface porosity/malformation with increasing acidification (seen primarily in the Jul 11 trial) described by the metrics introduced here reflect the results described in Kaplan et al. (2013). Squid statoliths are constructed through the growth of long, thin aragonite crystals from a core nucleation site within a protein matrix that directs the construction and expansion of the statolith (Radtke 1983). The aragonite crystals were long and thin, indicating a good calcification environment (high pH and aragonite saturation state) within the statocyst, suggesting that the disorientation of crystals and surficial degradation seen was instead an effect of decreased expression or activity of matrix proteins (Cohen and Holcomb 2009). Tests of paralarval swimming behavior, run in parallel to these experiments, demonstrated impacts of acidification on the energetics of swimming (primarily speed and vertical stationing), but did not show impairment to the paralarvae's ability to orient themselves or any aberrant swimming behaviors under hypercapnia (Zakroff et al. 2018). Given reported, dramatic responses of cephalopod paralarvae swimming behavior when statoliths are severely malformed or absent and hair cells are malfunctioning, these data suggest that despite observed statolith degradation, statocyst function may not have been severely impaired (Colmers et al. 1984; Hanlon et al. 1989; Zakroff et al. 2018). Due to the limitations of the image-based analyses performed, only a surficial description of the hatchling statolith can be considered. In further studies, it would be worthwhile to examine deeper layers or the density of the statolith to see when during embryonic development construction is disrupted by external stress.

A broader squid context

Many of the previous studies of OA and squid showed repeated significant effects on an array of variables (Lacoue-Labarthe et al. 2011; Kaplan et al. 2013; Hu et al. 2014; Rosa et al. 2014a; Navarro et al. 2016). Here, we had trials that were affected by relatively high levels of pCO_2 and low pH_t, but also trials that were not. This suggests some resiliency or tolerance of these squid to OA, at least during embryonic development. Indeed, these animals are tolerant of the naturally high pH and low oxygen concentrations of the egg capsule (Long et al. 2016). These results align with the limited, variable impacts of OA seen in *D. opalescens* embryos and are not unexpected when

879 considering the relative pCO_2 tolerance seen in juveniles and adults of *D. pealeii* and *D. gigas* (Rosa
880 and Seibel 2008; Seibel 2015, 2016; Navarro et al. 2016; Birk et al. 2018). Upregulation of key
881 proton secretion pathways in response to dramatic acidification (pH 7.31) in *Sepioteuthis*
882 *lessioniana* embryos also reinforces the scope for pH regulation and OA tolerance in this group (Hu
883 et al. 2013). In squids, physiological resiliency to OA may be species-specific, influenced by parental
884 environments, and/or under the influence of other unknown factors. Importantly, behavioral
885 sensitivity to OA has been shown in adult *Idiosepius pygmaeus*, which, while not a teuthid squid,
886 highlights the potential for neurologically driven impacts on these organisms that were not
887 examined here (Spady et al. 2014).

888 It has been suggested that marine invertebrates that produce egg capsules containing high
889 numbers of embryos have a substantial capacity for plasticity (Oyarzun and Strathmann 2011).
890 Cephalopods are broadly considered plastic organisms, altering their life history and population
891 structure under different environmental factors (Pecl et al. 2004; Pecl and Jackson 2008; Rosa et al.
892 2014b). Reproductive strategy and investment are also suggested to be highly plastic in
893 cephalopods, and are likely influenced by parental environment (Pecl and Moltschaniwskyj 2006;
894 Guerra et al. 2010; Robin et al. 2014). The dynamic variability in patterns of response to
895 acidification across metrics and trials demonstrated here might be a product of this squid's high
896 fecundity and patent plasticity.

897 As indicated by the potential relationship between our metrics and egg number, the
898 variability between culturing cups may act as an extension of variability between egg capsules.
899 Variability in the offspring of a single maternal clutch has been noted in the statoliths and DML of *S.*
900 *lessioniana* (Ikeda et al. 1999). Notable egg capsule variability has also been described in *D.*
901 *opalescens*, particularly in terms of statolith elemental composition (Navarro et al. 2014, 2016). In
902 the Aug 7 trial, variability between cups represented a very basic means of differentiating
903 parentage, maternity in particular, with embryos from tank B having slightly larger paralarvae with
904 slightly greater yolk (from a greater number of eggs per capsule) than tank A. Squid are not known
905 to maintain reserves of energy, not even for reproduction. Investment in reproduction primarily
906 depends on the tradeoff between overall somatic growth and the development of the reproductive
907 organs (Pecl and Moltschaniwskyj 2006). Production of eggs is fueled by energy captured through
908 feeding and so fecundity is linked with adult mantle length, as size acts as an indicator of both
909 energy intake potential and prey capture success (Boyle et al. 1995; Collins et al. 1995). While
910 degradation of maternal investment in successive clutches has been demonstrated in some

911 cephalopods, how a female squid distributes available energy among eggs and between egg
912 capsules of a single clutch is not well described to our knowledge, particularly among the multiple
913 egg per capsule squids (Steer et al. 2004).

914 Variation in offspring sensitivity to OA due to parental conditioning and epigenetics has
915 been described in fishes, often relating to seasonal variation in the population (Miller et al. 2012;
916 Murray et al. 2014; Schunter et al. 2016, 2018). Seasonal effects on sensitivity to OA have also been
917 described in *L. vulgaris*, with winter stock proving more resistant to both acidification and warming
918 (Rosa et al. 2014a). The distinctly different response patterns seen, across all metrics, between
919 trials suggests that some form of higher scale variability is occurring within the *D. pealeii* sensitivity
920 to OA stress. *Doryteuthis pealeii* has a roughly described, more anecdotally/locally acknowledged,
921 succession of size classes, which may be cohorts, across its breeding season (Arnold et al. 1974;
922 Mesnil 1977). Since these population dynamics are not well discriminated, a single year's sampling
923 is not substantive enough to determine whether the variation seen between trials represents a
924 consistent effect of seasonality/cohort on sensitivity to acidification stress. Further work would
925 require more replications over the course of the breeding season to parse out this variability.

926 In an organism as dynamic and complex as *D. pealeii* there are multiple scales of variability
927 to consider in assessing a physiological response to a stressor. This experiment served to highlight
928 small-scale variabilities: those between individuals, cups, days, and trials. These results also
929 highlight the importance of repetition and replication in organismal climate change response
930 studies, particularly with organisms that have a high potential for plasticity. As evidenced here,
931 neither data from a single trial nor data compiled across trials completely represented the scope of
932 this animal's sensitivity and tolerance to acidification (Fig. 3, Fig. 8). Further, dynamics of life
933 history must be considered in sampling, as parsing the data across days of hatching demonstrated.
934 At several points, across trials, had only certain days been sampled or only integrated data across
935 days been reported, the full dynamics of the stress response would not have been revealed (Fig. 4).
936 Investigation into sources of variability such as culture cup (which may relate to a previously
937 undescribed relationship with egg number) served to emphasize aspects of reproductive and
938 population biology that are still not well understood in this taxon and help guide needed future
939 work. Examination of data at all of these scales is valuable, although each may have its own utility,
940 but it is particularly worthwhile to examine these complex, frankly messy, systems as a whole as we
941 attempt to understand and predict how these organisms will fare in a rapidly changing ocean.

942

Compliance with Ethical Standards

Research involving human participants and/or animals

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

This material was based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1122374 to CZ. This project was funded by National Science Foundation Grant No. 1220034 to TAM.

Acknowledgments

We'd like to thank D. Remsen, the MBL Marine Resources Center staff, and MBL *Gemma* crew for their help acquiring squid. R. Galat and WHOI facilities staff provided system support. D. McCorkle, KYK Chan, and M. White provided guidance and insight on the acidification system and water quality monitoring. A. Solow provided statistics advice. We thank L. Kerr and the MBL Central Microscopy Facility for their aid with the SEM. We greatly appreciate E. Bonk, S. Zacarias, M. Lee, and A. Schlunk for their outstanding advice and assistance with this experiment. Thanks also to editors and anonymous reviewers for their constructive feedback on this manuscript.

969
970
971
972
973
974
975
976
977
978
979

980 **References**

981 Anthony KRN, Kline DI, Diaz-Pulido G, Dove S, Hoegh-Guldberg O (2008) Ocean acidification causes
982 bleaching and productivity loss in coral reef builders. Proc Natl Acad Sci 105:17442–17446.
983 doi: 10.1073/pnas.0804478105

984 Arnold JM, Summers WC, Gilbert DL, Manalis RS, Daw NW, Lasek RJ (1974) A guide to laboratory
985 use of the squid Loligo pealei. Marine Biological Laboratory, Woods Hole, MA

986 Bartol IK, Krueger PS, Thompson JT, Stewart WJ (2008) Swimming dynamics and propulsive
987 efficiency of squids throughout ontogeny. Integr Comp Biol 48:720–733. doi:
988 10.1093/icb/icn043

989 Baumann H, Wallace RB, Tagliaferri T, Gobler CJ (2014) Large Natural pH, CO2 and O2 Fluctuations
990 in a Temperate Tidal Salt Marsh on Diel, Seasonal, and Interannual Time Scales. Estuaries and
991 Coasts. doi: 10.1007/s12237-014-9800-y

992 Birk MA, McLean EL, Seibel BA (2018) Ocean acidification does not limit squid metabolism via
993 blood oxygen supply. J Exp Biol jeb.187443. doi: 10.1242/jeb.187443

994 Bonhomme V, Picq S, Gaucherel C, Claude J (2013) Momocs: outline analysis using R. J Stat Softw
995 56:1–24. doi: 10.18637/jss.v056.i13

996 Boyle PR, Pierce GJ, Hastie LC (1995) Flexible reproductive strategies in the squid *Loligo forbesi*.
997 Mar Biol 121:501–508.

998 Buresch KC, Maxwell MR, Cox MR, Hanlon RT (2009) Temporal dynamics of mating and paternity in
999 the squid *Loligo pealeii*. Mar Ecol Prog Ser 387:197–203. doi: 10.3354/meps08052

1000 Byrne M (2011) Impact of ocean warming and ocean acidification on marine invertebrate life
1001 history stages: Vulnerabilities and potential for persistence in a changing ocean. Ocean Mar
1002 Biol Annu Rev 49:1–42. doi: doi:10.1016/j.marenvres.2011.10.00

1003 Caldeira K, Wickett ME (2003) Oceanography: anthropogenic carbon and ocean pH. Nature
1004 425:365. doi: 10.1038/425365a

1005 Clayton TD, Byrne RH (1993) Spectrophotometric seawater pH measurements: total hydrogen
1006 results. Deep Res 40:2115–2129.

1007 Cohen AL, Holcomb M (2009) Why corals care about ocean acidification: Uncovering the
1008 mechanism. Oceanography 22:118–127. doi: 10.5670/oceanog.2009.102

1009 Collins MA, Burnell GM, Rodhouse PG (1995) Reproductive strategies of male and female *Loligo*
1010 *forbesi* (Cephalopoda: Loliginidae). J Mar Biol Assoc UK 75:621–634.

1011 Colmers WF, Hixon RF, Hanlon RT, Forsythe JW, Ackerson M V., Wiederhold ML, Hulet WH (1984)
1012 Spinner cephalopods: defects of statocyst suprastructures in an invertebrate analogue of the
1013 vestibular apparatus. Cell Tissue Res. doi: 10.1007/BF00217217

1014 Dickson AG (1990) Standard potential of the reaction: AgCl(s) + (1/2)H_2(g) = Ag(s) + HCl(aq), and
1015 and the standard acidity constant of the ion HSO_4^- in synthetic sea water from 273.15 to
1016 318.15 K. J Chem Thermodyn 22:113–127. doi: 10.1016/0021-9614(90)90074-Z

1017 Dickson AG, Sabine CL, Christian JR (2007) Guide to best practices for ocean CO_2 measurements.
1018 PICES Spec Publ 3:p191. doi: 10.1159/000331784

1019 Doney SC, Fabry VJ, Feely RA, Kleypas JA (2009) Ocean acidification: the other CO_2 problem. Ann
1020 Rev Mar Sci 1:169–192. doi: 10.1146/annurev.marine.010908.163834

1021 Fabry VJ, Seibel BA, Feely RA, Orr JC (2008) Impacts of ocean acidification on marine fauna and
1022 ecosystem processes. ICES J Mar Sci 65:414. doi: 10.1093/icesjms/fsn048

1023 Gallager SM, Mann R, Sasaki GC (1986) Lipid as an index of growth and viability in three species of
1024 bivalve larvae. Aquaculture 56:81–103. doi: 10.1016/0044-8486(86)90020-7

1025 Gazeau F, Parker LM, Comeau S, Gattuso J-PP, O'Connor WA, Martin S, Pörtner H-O, Ross PM (2013)
1026 Impacts of ocean acidification on marine shelled molluscs. Mar Biol 160:2207–2245. doi:
1027 10.1007/s00227-013-2219-3

1028 Gledhill DK, White MM, Salisbury J, Thomas H, Misna I, Liebman M, Mook B, Grear J, Candelmo AC,
1029 Chambers RC, Gobler CJ, Hunt CW, King AL, Price NN, Signorini SR, Stancioff E, Stymiest C,

Wahle RA, Waller JD, Rebuck ND, Wang ZA, Capson TL, Morrison JR, Cooley SR, Doney SC (2015) Ocean and coastal acidification off New England and Nova Scotia. Oceanography 28:182–197. doi: http://dx.doi.org/10.5670/oceanog.2015.41

Gray CL (1992) Long-finned Squid (*Loligo pealei*) Species Profile. In: Current Report: The Narragansett Bay Project NBP-92-106. pp 1–54

Guerra Á, Allcock L, Pereira J (2010) Cephalopod life history, ecology and fisheries: An introduction. Fish Res 106:117–124. doi: 10.1016/j.fishres.2010.09.002

Gutowska MA, Melzner F (2009) Abiotic conditions in cephalopod (*Sepia officinalis*) eggs: Embryonic development at low pH and high pCO_2. Mar Biol 156:515–519. doi: 10.1007/s00227-008-1096-7

Haigh R, Ianson D, Holt CA, Neate HE, Edwards AM (2015) Effects of ocean acidification on temperate coastal marine ecosystems and fisheries in the northeast Pacific. PLoS One 10:e0117533. doi: 10.1371/journal.pone.0117533

Hanlon R, Bidwell J, Tait R (1989) Strontium is required for statolith development and thus normal swimming behaviour of hatchling cephalopods. J Exp Biol 141:187–195.

Hanlon RT, Messenger JB (1998) Cephalopod Behaviour. Cambridge University Press, Cambridge, UK

Hofmann GE, Todgham AE (2010) Living in the now: Physiological mechanisms to tolerate a rapidly changing environment. Annu Rev Physiol 72:127–145. doi: 10.1146/annurev-physiol-021909-135900

Hu MY, Sucre E, Charmantier-Daures M, Charmantier G, Lucassen M, Himmerkus N, Melzner F (2010) Localization of ion-regulatory epithelia in embryos and hatchlings of two cephalopods. Cell Tissue Res 339:571–583.

Hu MY, Tseng Y-C, Stumpp M, Gutowska MA, Kiko R, Lucassen M, Melzner F (2011a) Elevated seawater pCO_2 differentially affects branchial acid-base transporters over the course of development in the cephalopod *Sepia officinalis*. Am J Physiol Regul Integr Comp Physiol 300:R1100–R1114. doi: 10.1152/ajpregu.00653.2010

Hu MY, Tseng Y-C, Lin L-Y, Chen P-Y, Charmantier-Daures M, Hwang P-P, Melzner F (2011b) New insights into ion regulation of cephalopod molluscs: a role of epidermal ionocytes in acid-base regulation during embryogenesis. AJP Regul Integr Comp Physiol 301:R1700–R1709. doi: 10.1152/ajpregu.00107.2011

Hu MY, Lee J-R, Lin L-Y, Shih T-H, Stumpp M, Lee M-F, Hwang P-P, Tseng Y-C (2013) Development in a naturally acidified environment: Na^+/H^+-exchanger 3-based proton secretion leads to CO_2 tolerance in cephalopod embryos. Front Zool 10:51. doi: 10.1186/1742-9994-10-51

Hu MY, Guh Y-J, Stumpp M, Lee J-R, Chen R-D, Sung P-H, Chen Y-C, Hwang P-P, Tseng Y-C (2014)

Branchial NH_4^+-dependent acid–base transport mechanisms and energy metabolism of squid (*Sepioteuthis lessoniana*) affected by seawater acidification. Front Zool 11:55. doi: 10.1186/s12983-014-0055-z

Ikeda Y, Wada Y, Arai N, Sakamoto W (1999) Note on size variation of body and statoliths in the oval squid *Sepioteuthis lessoniana* hatchlings. J Mar Biol Assoc UK 79:757–759. doi: 10.1017/S0025315498000939

Jacobson LD (2005) Longfin inshore squid, Loligo pealeii, life history and habitat characteristics. In: NOAA Technical Memorandum NMFS-NE-193. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA, pp 1–42

Kaplan MB, Mooney TA, McCorkle DC, Cohen AL (2013) Adverse effects of ocean acidification on early development of squid (*Doryteuthis pealeii*). PLoS One 8:e63714. doi: 10.1371/journal.pone.0063714

Lacoue-Labarthe T, Réveillac E, Oberhänsli F, Teyssié JL, Jeffree R, Gattuso JP (2011) Effects of ocean acidification on trace element accumulation in the early-life stages of squid *Loligo vulgaris*. Aquat Toxicol 105:166–176. doi: 10.1016/j.aquatox.2011.05.021

Langsrud Ø (2003) ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. Stat Comput 13:163–167. doi: 10.1023/A:1023260610025

Laptikhovsky V, Nikolaeva S, Rogov M (2018) Cephalopod embryonic shells as a tool to reconstruct reproductive strategies in extinct taxa. Biol Rev 93:270–283. doi: 10.1111/brv.12341

Long MH, Mooney TA, Zakroff C (2016) Extreme low oxygen and decreased pH conditions naturally occur within developing squid egg capsules. Mar Ecol Prog Ser 550:111–119. doi: 10.3354/meps11737

Martins RS, Roberts MJ, Chang N, Verley P, Moloney CL, Vidal E a G (2010a) Effect of yolk utilization on the specific gravity of chokka squid (*Loligo reynaudii*) paralarvae: Implications for dispersal on the Agulhas Bank, South Africa. ICES J Mar Sci 67:1323–1335. doi: 10.1093/icesjms/fsq098

Martins RS, Roberts MJ, Vidal ÉAG, Moloney CL (2010b) Effects of temperature on yolk utilization by chokka squid (*Loligo reynaudii* d'Orbigny, 1839) paralarvae. J Exp Mar Bio Ecol 386:19–26. doi: 10.1016/j.jembe.2010.02.014

Maxwell MR, Hanlon RT (2000) Female reproductive output in the squid *Loligo pealeii*: Multiple egg clutches and implications for a spawning strategy. Mar Ecol Prog Ser 199:159–170. doi: 10.3354/meps199159

McMahon JJ, Summers WC (1971) Temperature effects on the developmental rate of squid (*Loligo pealei*) embryos. Biol Bull 141:561–567.

Mehrbach C, Culberson CH, Hawley JE, Pytkowicz RM (1973) Measurement of the apparent

dissociation constants of carbonic acid in seawater at atmospheric pressure. Limnol Oceanogr 18:897–907. doi: 10.4319/lo.1973.18.6.0897

Mesnil B (1977) Growth and Life Cycle of Squid, Loligo pealei and Illex illecebrosus, from the Northwest Atlantic. In: International Commission for the Northwest Atlantic Fisheries Selected Papers. pp 55–69

Michaelidis B, Ouzounis C, Paleras A, Pörtner H-O (2005) Effects of long-term moderate hypercapnia on acid-base balance and growth rate in marine mussels *Mytilus galloprovincialis*. Mar Ecol Prog Ser 293:109–118. doi: 10.3354/meps293109

Miller GM, Watson S-A, Donelson JM, McCormick MI, Munday PL (2012) Parental environment mediates impacts of increased carbon dioxide on a coral reef fish. Nat Clim Chang 2:858–861. doi: 10.1038/nclimate1599

Murray CS, Malvezzi A, Gobler CJ, Baumann H (2014) Offspring sensitivity to ocean acidification changes seasonally in a coastal marine fish. Mar Ecol Prog Ser 504:1–11. doi: 10.3354/meps10791

Navarro MO, Bockmon EE, Frieder CA, Gonzalez JP, Levin LA (2014) Environmental pH, O_2 and capsular effects on the geochemical composition of statoliths of embryonic squid *Doryteuthis opalescens*. Water 2233–2254. doi: 10.3390/w6082233

Navarro MO, Kwan GT, Batalov O, Choi CY, Pierce NT, Levin LA (2016) Development of embryonic market squid, Doryteuthis opalescens, under chronic exposure to low environmental pH and [O2]. PLoS One 11:e0167461. doi: 10.1371/journal.pone.0167461

Navarro MO, Parnell PE, Levin LA (2018) Essential market squid (*Doryteuthis opalescens*) Embryo Habitat: A Baseline for Anticipated Ocean Climate Change. J Shellfish Res 37:601–614. doi: 10.2983/035.037.0313

NOAA (2019) Squid, Mackerel, and Butterfish Quota Monitoring Page. In: NOAA Fish. - Gt. Atl. Reg. https://www.greateratlantic.fisheries.noaa.gov/aps/monitoring/longfinsquid.html. Accessed 16 Mar 2019

Oyarzun FX, Strathmann RR (2011) Plasticity of hatching and the duration of planktonic development in marine invertebrates. Integr Comp Biol 51:81–90. doi: 10.1093/icb/icr009

Pecl GT, Jackson GD (2008) The potential impacts of climate change on inshore squid: Biology, ecology and fisheries. Rev Fish Biol Fish 18:373–385. doi: 10.1007/s11160-007-9077-3

Pecl GT, Moltschaniwskyj NA (2006) Life history of a short-lived squid (*Sepioteuthis australis*): resource allocation as a function of size, growth, maturation, and hatching season. ICES J Mar Sci 63:995–1004. doi: 10.1016/j.icesjms.2006.04.007

Pecl GT, Moltschaniwskyj NA, Tracey SR, Jordan AR (2004) Inter-annual plasticity of squid life history and population structure: Ecological and management implications. Oecologia

139:515–524. doi: 10.1007/s00442-004-1537-z

Pierrot D, Lewis E, Wallace DWR (2006) MS Excel program developed for CO2 system calculations. In: ORNL/CDIAC-105a. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, pp 1–17

Pörtner H-O, Reipschlager A, Heisler N (1998) Acid-base regulation, metabolism and energetics in *Sipunculus nudus* as a function of ambient carbon dioxide level. J Exp Biol 201:43–55.

Radtke RL (1983) Chemical and structural characteristics of statoliths from the short-finned squid *Illex illecebrosus*. Mar Biol 76:47–54. doi: 10.1007/BF00393054

Robin JP, Roberts M, Zeidberg L, Bloor I, Rodriguez A, Briceño F, Downey N, Mascaró M, Navarro M, Guerra A, Hofmeister J, Barcellos DD, Lourenço SAP, Roper CFE, Moltschaniwskyj NA, Green CP, Mather J (2014) Transitions during cephalopod life history: The role of habitat, environment, functional morphology and behaviour. In: Vidal EAG (ed) Advances in Cephalopod Science: Biology, Ecology, Cultivation and Fisheries. Academic Press, Cambridge, MA, pp 361–437

Rosa R, Seibel BA (2008) Synergistic effects of climate-related variables suggest future physiological impairment in a top oceanic predator. Proc Natl Acad Sci 105:20776–20780. doi: 10.1073/pnas.0806886105

Rosa R, Trübenbach K, Repolho T, Pimentel M, Faleiro F, Boavida-Portugal J, Baptista M, Lopes VM, Dionísio G, Leal MC, Calado R, Pörtner HO (2013) Lower hypoxia thresholds of cuttlefish early life stages living in a warm acidified ocean. Proc R Soc B Biol Sci 280:20131695. doi: 10.1098/rspb.2013.1695

Rosa R, Trübenbach K, Pimentel MS, Boavida-Portugal J, Faleiro F, Baptista M, Dionísio G, Calado R, Pörtner HO, Repolho T (2014a) Differential impacts of ocean acidification and warming on winter and summer progeny of a coastal squid (Loligo vulgaris). J Exp Biol 217:518–25. doi: 10.1242/jeb.096081

Rosa R, O'Dor R, Pierce G (2014b) Myopsid Squids. Nova Science Publishers, Inc, New York, NY

Schunter C, Welch MJ, Ryu T, Zhang H, Berumen ML, Nilsson GE, Munday PL, Ravasi T (2016) Molecular signatures of transgenerational response to ocean acidification in a species of reef fish. Nat Clim Chang 6:1014–1018. doi: 10.1038/nclimate3087

Schunter C, Welch MJ, Nilsson GE, Rummer JL, Munday PL, Ravasi T (2018) An interplay between plasticity and parental phenotype determines impacts of ocean acidification on a reef fish. Nat Ecol Evol 2:334–342. doi: 10.1038/s41559-017-0428-8

Seibel BA (2015) Environmental physiology of the jumbo squid, *Dosidicus gigas* (d 'Orbigny, 1835) (Cephalopoda : Ommastrephidae): Implications for changing climate. Am Malacol Bull 33:1–13.

1170 Seibel BA (2016) Cephalopod susceptibility to asphyxiation via ocean incalescence, deoxygenation,
1171 and acidification. Physiology 31:418–429. doi: 10.1152/physiol.00061.2015

1172 Spady BL, Watson S, Chase TJ, Munday PL (2014) Projected near-future CO_2 levels increase activity
1173 and alter defensive behaviours in the tropical squid *Idiosepius pygmaeus*. Biol Open 3:1063–
1174 70. doi: 10.1242/bio.20149894

1175 Steer M, Moltschaniwskyj N, Nichols D, Miller M (2004) The role of temperature and maternal
1176 ration in embryo survival: Using the dumpling squid *Euprymna tasmanica* as a model. J Exp
1177 Mar Bio Ecol 307:73–89. doi: 10.1016/j.jembe.2004.01.017

1178 Steer MA, Pecl GT, Moltschaniwskyj NA (2003) Are bigger calamary *Sepioteuthis australis* hatchlings
1179 more likely to survive? A study based on statolith dimensions. Mar Ecol Prog Ser 261:175–182.
1180 doi: 10.3354/meps261175

1181 Vidal EAG, Haimovici M (1998) Feeding and the possible role of the proboscis and mucus cover in
1182 the ingestion of microorganism by rhynchoteuthion paralarvae (Cephalopoda:
1183 Ommastrephidae). Bull Mar Sci 63:305–316.

1184 Vidal EAG, DiMarco FP, Wormuth JH, Lee PG (2002a) Influence of temperature and food availability
1185 on survival, growth and yolk utilization in hatchling squid. Bull Mar Sci 71:915–931.

1186 Vidal EAG, DiMarco FP, Wormuth JH, Lee PG (2002b) Optimizing rearing conditions of hatchling
1187 loliginid squid. Mar Biol 140:117–127. doi: 10.1007/s002270100683

1188 Villanueva R, Quintana D, Petroni G, Bozzano A (2011) Factors influencing the embryonic
1189 development and hatchling size of the oceanic squid Illex coindetii following in vitro
1190 fertilization. J Exp Mar Bio Ecol 407:54–62. doi: 10.1016/j.jembe.2011.07.012

1191 Wang ZA, Wanninkhof R, Cai W-J, Byrne RH, Hu X, Peng T-H, Huang W-J (2013) The marine
1192 inorganic carbon system along the Gulf of Mexico and Atlantic coasts of the United States :
1193 Insights from a transregional coastal carbon study. Limnol Oceanogr 58:325–342. doi:
1194 10.4319/lo.2013.58.1.0325

1195 White MM, McCorkle DC, Mullineaux LS, Cohen AL (2013) Early exposure of bay scallops
1196 (*Argopecten irradians*) to high CO_2 causes a decrease in larval shell growth. PLoS One 8:2–9.
1197 doi: 10.1371/journal.pone.0061065

1198 York CA, Bartol IK (2016) Anti-predator behavior of squid throughout ontogeny. J Exp Mar Bio Ecol
1199 480:26–35. doi: 10.1016/j.jembe.2016.03.011

1200 Zakroff C, Mooney TA, Wirth C (2018) Ocean acidification responses in paralarval squid swimming
1201 behavior using a novel 3D tracking system. Hydrobiologia 808:83–106. doi: 10.1007/s10750-
1202 017-3342-9

1203

List of Abbreviations

2D - Two-dimensional

DML - Dorsal mantle length

ESL - Environmental Systems Laboratory

KW - Kruskal-Wallis test

LR - Linear Regression

MBL - Marine Biological Laboratory

OA - Ocean acidification

SEM - Scanning electron microscopy

YV - Yolk Volume

Fig. 1

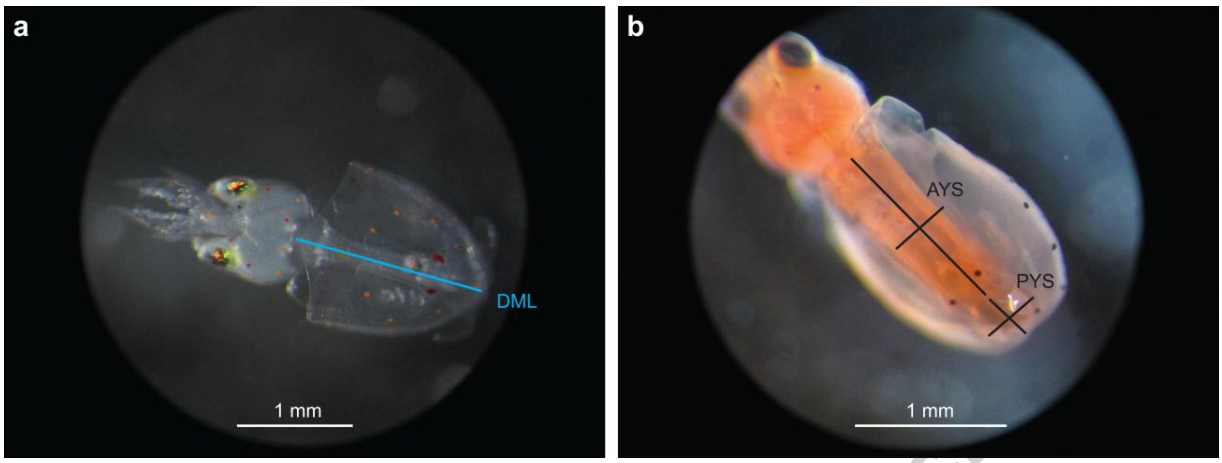
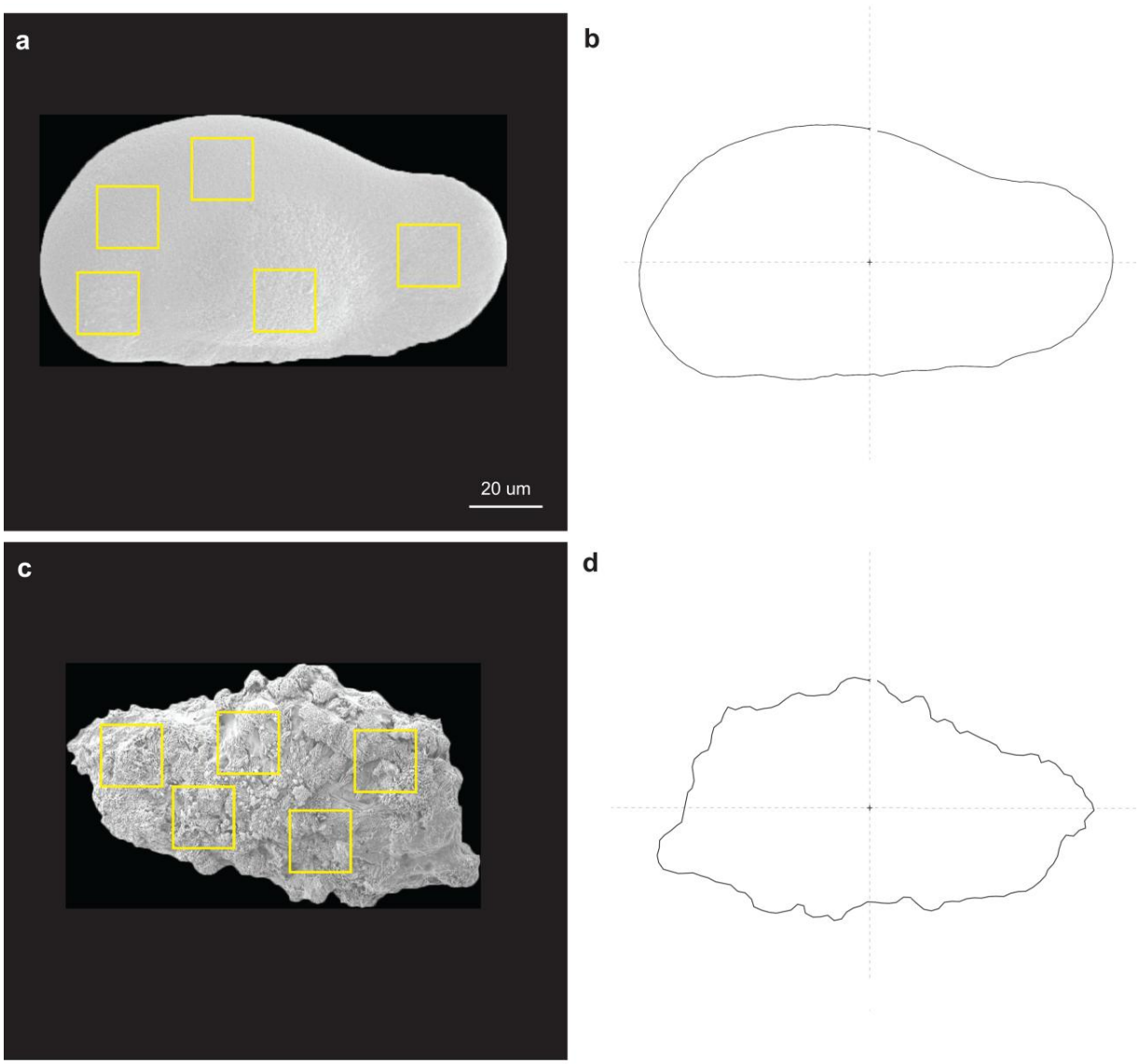


Fig. 1

Doryteuthis pealeii paralarvae imaged for measurements of dorsal mantle length and yolk sac volume. **a** An anaesthetized paralarva photographed for measurement of its dorsal mantle length (DML, superimposed cyan line). **b** A preserved paralarva stained with oil red O photographed for measurement of its yolk sac volume. Length and width (superimposed black lines) of the anterior yolk sac (AYS) and posterior yolk (PYS) were measured to calculate total yolk volume. Scale bars are unique to each image, both representing 1 mm. Photos by CZ

Fig. 2



1243

1244

1245 **Fig. 2**

1246 Statoliths extracted from control and high acidification exposures measured for size, shape, and
1247 surface variability. **a** A statolith from the 400 ppm CO_2 treatment with the MATLAB surface analysis
1248 squares superimposed in yellow and **b** its outline produced in the R Momocs package. **c** A statolith
1249 from the 2200 ppm treatment with analysis squares superimposed and **d** its outline from Momocs.
1250 All images are to the same 20 μm scale, shown in **a**

1251 **Fig. 3**

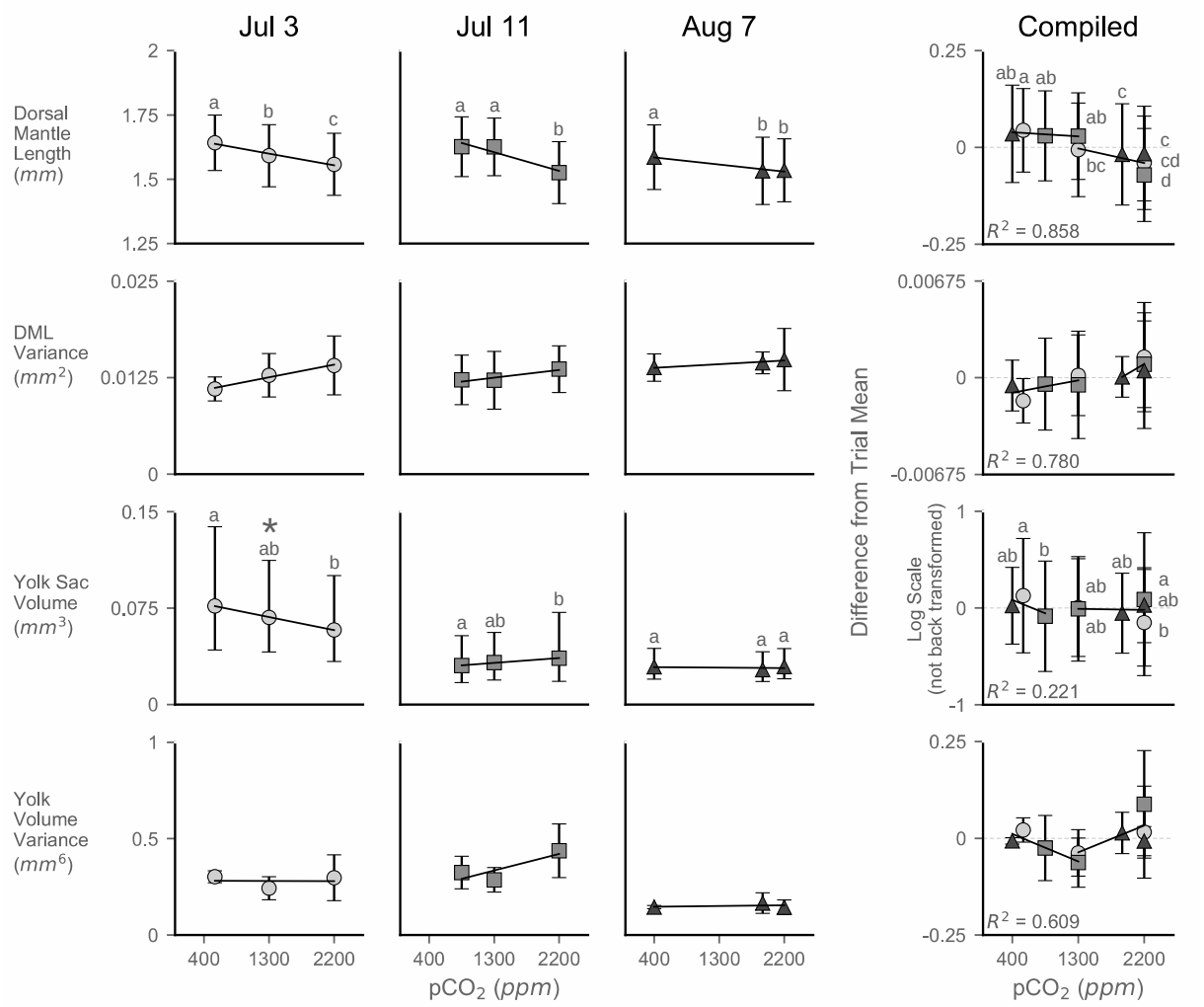
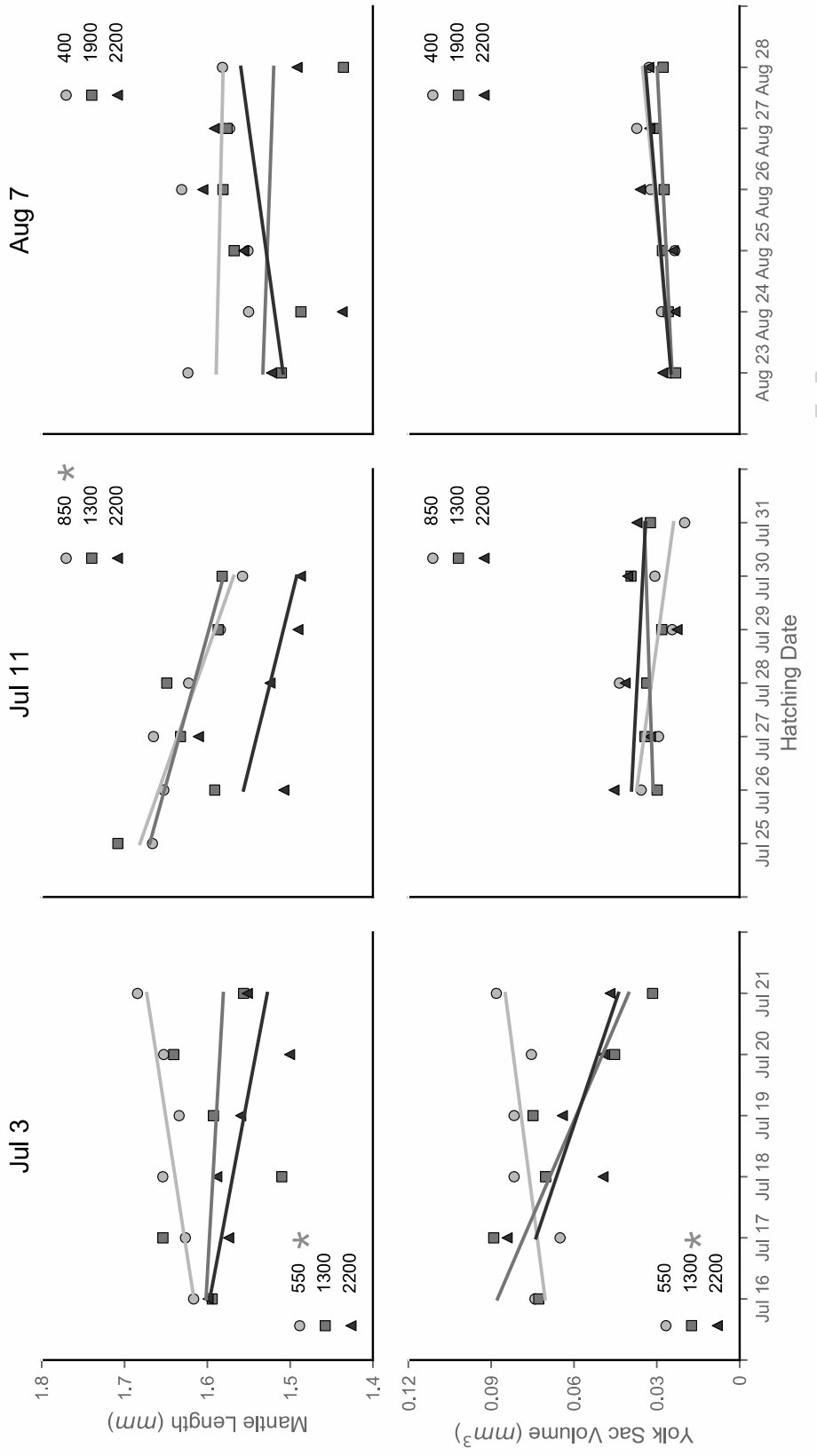


Fig. 3

Dorsal mantle length, yolk sac volume, and respective variances of paralarvae exposed to a range of pCO_2 treatments. Data are presented separated by trial (demarcated by egg capsule laying date) compiled across cups and hatching days for each pCO_2 treatment (metric n's in Table 1, variance n = 3 cups per treatment per trial). The Compiled plot depicts the data from all trials normalized by taking sample values and subtracting its respective trial mean. Differences in log transformed yolk sac volume data are not back transformed. Symbols represent means, with shape and color corresponding to trial. Error bars represent one standard deviation. Letters demarcate statistical groupings from a Tukey's HSD. Trend lines in trial data depict linear regressions; significance is marked with an asterisk ($P < 0.05$). Models of best fit from piecewise regressions are presented on compiled data with corresponding R^2

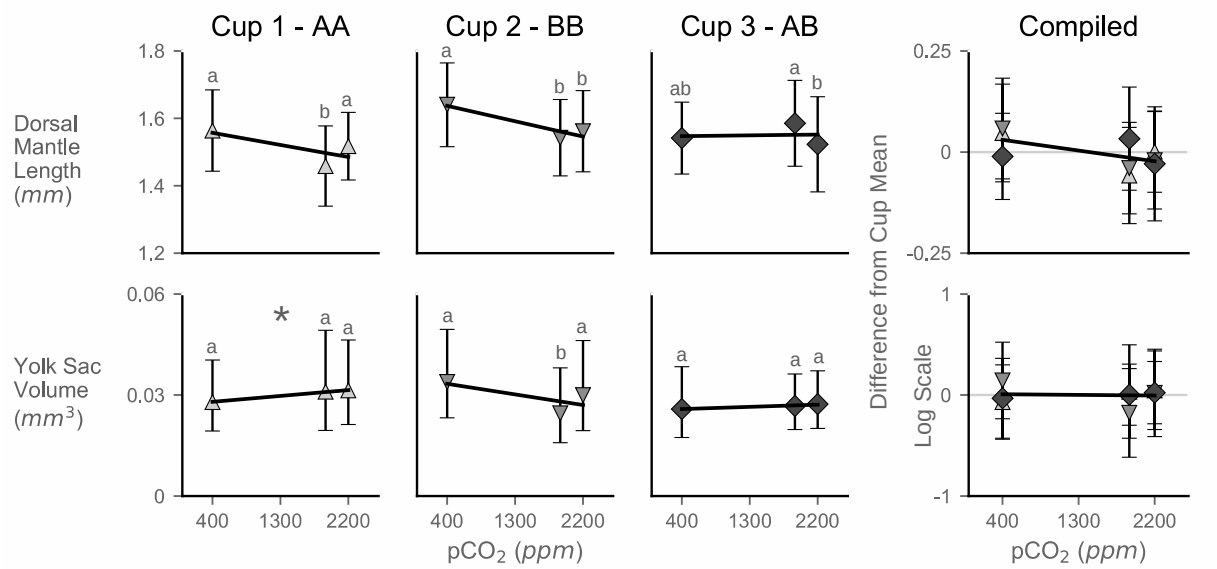


1264

Fig. 4

Mean dorsal mantle length and yolk sac volume (back transform of logarithmic mean) of paralarvae across sampled hatching days. Measurements for the Jul 3, Jul 11, and Aug 7 trials are compiled across cups and presented by CO_2 treatment; n ~ 30 (~10 per experimental cup) paralarvae per symbol. Symbols represent means, with shape and color corresponding to pCO_2 treatment (ppm). Error is not shown for visual clarity. Linear regressions are colored corresponding to their pCO_2 treatment; significance is marked with an asterisk next to pCO_2 treatment in the legend

1292 **Fig. 5**



1293

1294 **Fig. 5**

1295 Dorsal mantle length and yolk sac volume (back transformed from logarithmic data) of Aug 7 trial
1296 paralarvae separated by culture cup. Cups in the Aug 7 trial each contained two egg capsules sorted
1297 from two separate adult squid tanks, tank A and tank B (Cup 1 = AA, Cup 2 = BB, and Cup 3 = AB).
1298 The Compiled plot depicts the data from all cups normalized by taking sample values and
1299 subtracting its respective cup mean. Differences in log transformed yolk data are not back
1300 transformed. Symbols represent means, with shape and color corresponding to cup. Error bars
1301 represent one standard deviation; n = ~53 paralarvae per symbol (~10 per day for 6 days, often
1302 fewer in the latter days of hatching). Letters demarcate statistical groupings from a Tukey's HSD.
1303 Lines depict linear regressions; significance is marked with an asterisk ($P < 0.05$)

1304

1305

1306

1307

1308

1309

1310 **Fig. 6**

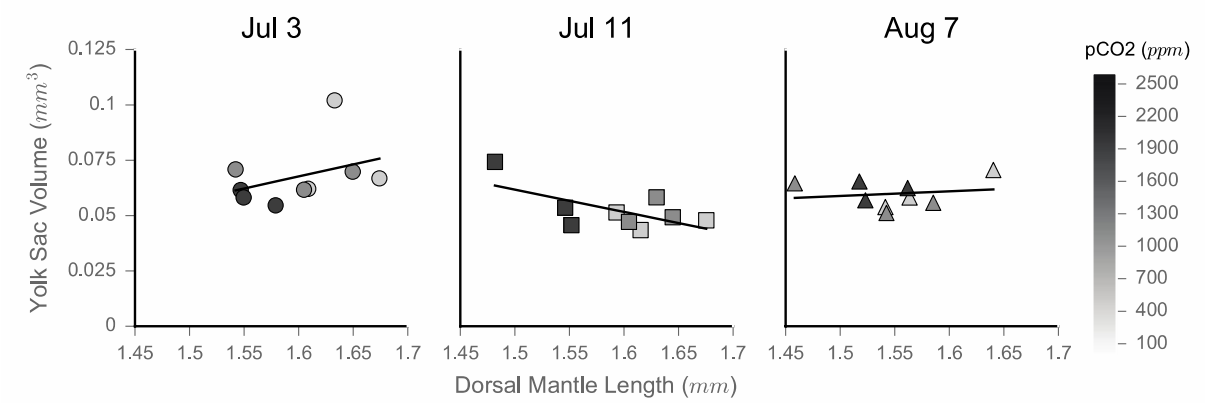
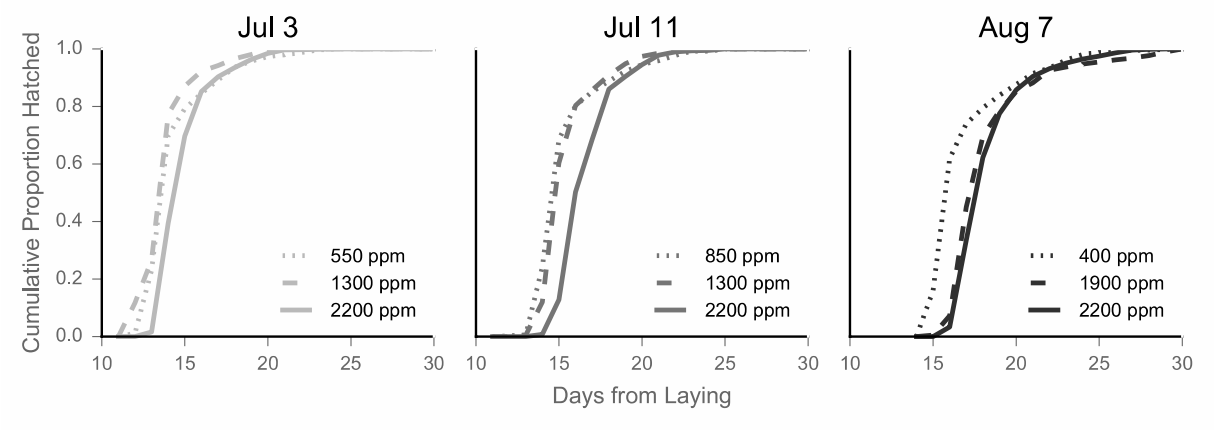


Fig. 6

Comparison of average yolk sac volume and average mantle length. Data are averaged for each culture cup and are presented separated by trial; n = 3 experimental cups per treatment per trial. Error bars for both axes are not depicted for visual clarity and to focus on trend lines. Symbols represent means, with shape corresponding to trial, and color corresponding to pCO_2 (color bar at right). Lines depict linear regressions; none were significant ($P < 0.05$)

1331 **Fig. 7**



1332

1333 **Fig. 7**

1334 Hatching time curves for each pCO_2 treatment. Hatching counts are plotted as the cumulative
1335 percent hatching per day to produce smooth curves. Data are plotted by trial, denoted by lay date
1336 (titles) and color; n = 3 experimental cups (with 2 egg capsules each) per treatment per trial. Error
1337 bars/shading not depicted for visual clarity of the curves. Line patterning demarcates pCO_2
1338 treatment, with lines becoming more solid with increasing acidification

1339

1340

1341

1342

1343

1344

1345

1346

1347

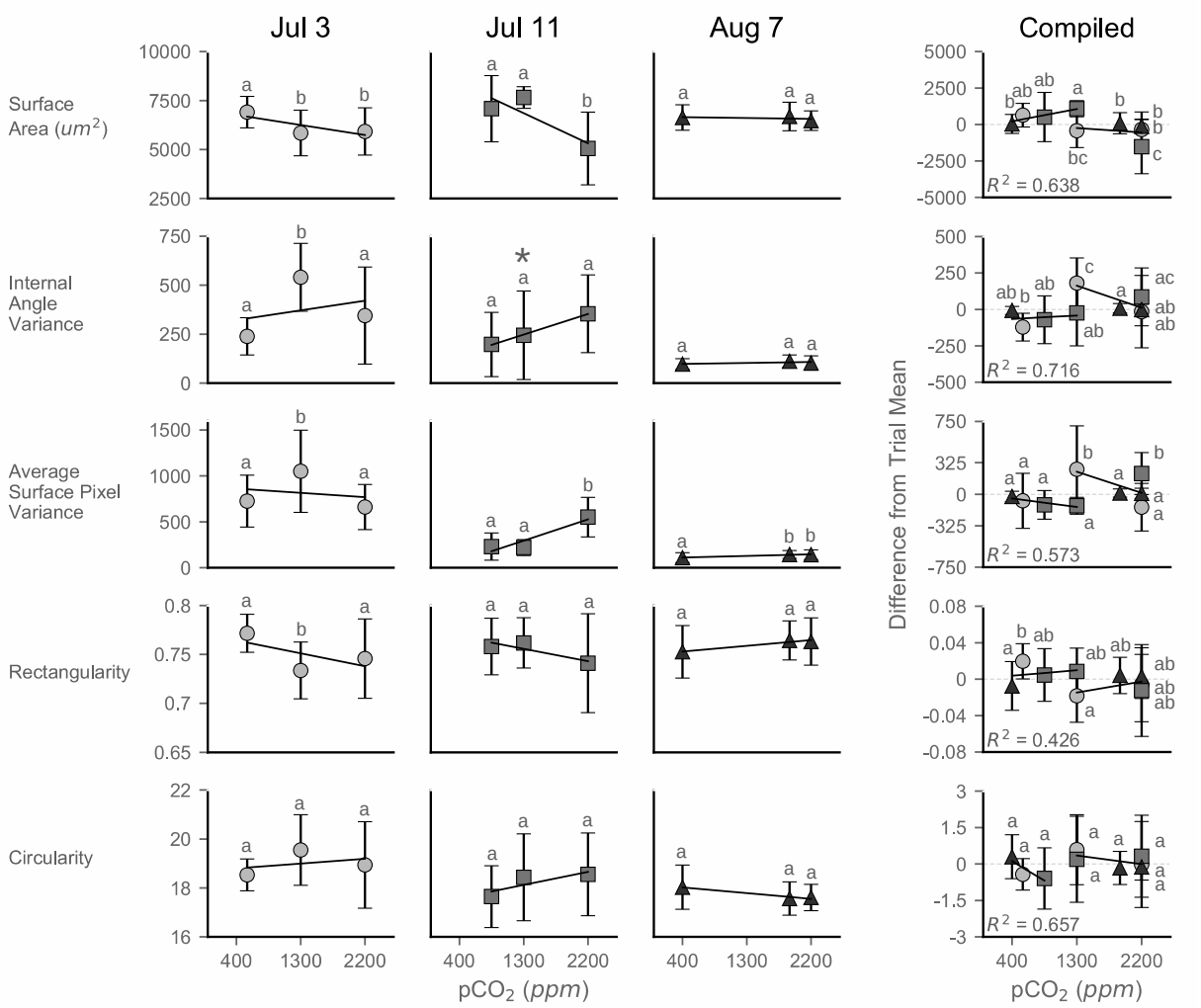
1348

1349

1350

1351

1352 **Fig. 8**



1353

1354 **Fig. 8**

1355 Statolith morphometrics across a range of pCO_2 treatments. Data are presented separated by trial
1356 (demarcated by egg capsule laying date) compiled across cups and hatching days for each pCO_2
1357 treatment. The Compiled plot depicts the data from all trials normalized by taking sample values
1358 and subtracting its respective trial mean (n's in Table 1). Models of best fit from piecewise
1359 regressions are presented on compiled data with corresponding R^2 values. Symbols represent
1360 means, with shape and color corresponding to trial. Error bars represent one standard deviation.
1361 Letters demarcate statistical groupings from a Dunn's test. Lines depict linear regressions;
1362 significance is marked with an asterisk ($P < 0.05$)

Table 1. Seawater chemistry measurements and number of samples taken for each treatment of each trial.

Laying Date	Treatment pCO_2 (ppm)	Temp (°C)	pH_{total}	Salinity	A_T (mmol $kgSW^{-1}$)	Ω_{Arag}	pCO_2 (ppm)	n		
								Mantle Length	Yolk Sac Volume	Statoliths
3-Jul	Ambient (550)	20.83 ± 0.19	7.88 (0.02)	31.40 (0.06)	2060.3 (12.5)	1.87 (0.08)	565.68 (43.90)	162	155	20
	1300	20.83 ± 0.19	7.54 (0.01)	31.41 (0.03)	2064.7 (6.6)	0.93 (0.04)	1350.51 (43.55)	145	121	15
	2200	20.83 ± 0.19	7.34 (0.01)	31.39 (0.05)	2064.9 (16.1)	0.60 (0.04)	2199.56 (173.47)	158	136	19
11-Jul	850	20.46 (0.03)	7.66 (0.01)	31.26 (0.13)	2042.1 (30.3)	1.17 (0.03)	987.43 (20.30)	172	160	13
	1300	20.46 (0.03)	7.54 (0.03)	31.29 (0.07)	2051.6 (10.2)	0.90 (0.03)	1351.67 (34.26)	175	166	15
	2200	20.46 (0.03)	7.31 (0.01)	31.25 (0.13)	2047.1 (21.2)	0.54 (0.02)	2380.50 (70.62)	144	169	15
7-Aug	400	19.97 (0.49)	7.93 (0.01)	31.51 (0.01)	2032.0 (20.0)	1.95 (0.00)	488.58 (10.50)	161	159	31
	1900	19.97 (0.49)	7.37 (0.00)	31.46 (0.02)	2015.8 (5.7)	0.59 (0.00)	2003.79 (12.84)	163	161	32
	2200	19.97 (0.49)	7.35 (0.01)	31.45 (0.02)	2028.1 (0.0)	0.56 (0.01)	2130.17 (40.31)	169	160	34

Seawater data are presented as means ± standard deviation, n = 3 samplings of the control cup for each treatment in each trial

1363

Table 2. Three-way Type II nested ANOVAs for compiled data and individual trials, for both mantle length and (log-transformed) yolk volume. Significant p values (α = 0.05) in bold.

	Mantle Length					Yolk Sac Volume				
Source	SS	df	*F*	*P*	Ω^2	SS	df	*F*	*P*	Ω^2
Compiled Data										
Trial	0.418	2	17.39	**<0.001**	0.013	0.002	2	1.478	0.229	0.000
Trial : pCO_2	0.562	15	3.121	0.075	0.013	0.003	15	0.243	0.983	-0.005
Trial : Date	$-2.27*10^{-9}$	51	$-3.703*10^{-9}$	1.000	-0.020	$1.4*10^{-5}$	51	$3.44*10^{-4}$	1.000	-0.023
Trial : pCO_2 : Date	9.599	255	3.135	0.077	0.214	0.429	255	2.164	0.142	0.136
Trial : pCO_2 : Cup	3.355	36	7.760	**<0.001**	0.096	0.238	36	8.501	**<0.001**	0.128
Residual	16.55	1378				1.023	1316			
Jul 3										
pCO_2	0.001	2	0.048	0.828	-0.003	$-4.836*10^{-12}$	2	$-1.615*10^{-9}$	1.000	0.000
Date	0.066	5	1.296	0.271	0.002	0.039	5	5.253	**0.001**	0.042
pCO_2 : Date	0.563	10	5.491	**<0.001**	0.070	0.169	10	11.13	**<0.001**	0.180
pCO_2 : Cup	0.576	6	9.363	**<0.001**	0.078	0.078	6	8.730	**<0.001**	0.084
pCO_2 : Date : Cup	1.150	30	3.739	**<0.001**	0.127	0.105	30	2.330	**<0.001**	0.111
Residual	4.256	415				0.548	366			
Jul 11										
pCO_2	0.162	2	7.473	**<0.001**	0.018	0.007	2	5.222	**0.006**	0.015
Date	0.908	5	16.80	**<0.001**	0.111	0.010	5	3.108	**0.009**	0.018
pCO_2 : Date	0.895	10	8.274	**<0.001**	0.103	0.013	10	1.990	**0.033**	0.017
pCO_2 : Cup	0.233	6	3.589	**0.002**	0.022	0.024	6	6.189	**<0.001**	0.054
pCO_2 : Date : Cup	0.669	30	2.061	**0.001**	0.045	0.038	30	1.860	**0.004**	0.044
Residual	4.759	440				0.290	441			
Aug 7										
pCO_2	0.326	2	13.81	**<0.001**	0.036	$4.81*10^{-4}$	2	1.670	0.190	0.002
Date	0.828	5	14.06	**<0.001**	0.091	0.005	5	7.608	**<0.001**	0.057
pCO_2 : Date	0.457	10	3.874	**<0.001**	0.040	0.002	10	1.665	0.095	0.012
pCO_2 : Cup	1.054	6	14.90	**<0.001**	0.116	0.005	6	5.590	**<0.001**	0.048
pCO_2 : Date : Cup	0.603	30	1.705	**0.013**	0.029	0.008	30	1.806	**0.007**	0.042
Residual	5.186	440				0.062	428			

1364