**Statistical Methods**

*Lethal Data*

To examine the effect of treatment on sea cucumber mortality, we ran an ANOVA test from the {stats} R package {REF}. We used the fitDist function from the {gamlss} R package to select the appropriate distribution and used forward selection to create a logistic regression model with pooping status, weight, and initial droop and squeeze scores as explanatory variables in the full model. We ran the model to investigate the significance of the selected explanatory variable(s).

*Sublethal Data*

We first examined how cucumber stiffness, indicated by antipredator response and structure maintenance, varied across the experiment using stacked bar charts. We investigated the degree of correlation between antipredator response and structure maintenance stiffness using spearman’s correlation coefficient, but modelled them as separate responses because we assume they represent different types of stiffening behaviour. We then used the *clmm* function from package ordinal (CITE) to conduct ordered logistic regression, also known as ordinal regression. For both response variables, we constructed full models with the following predictor variables: treatment (12C, 17C, 22C) as a categorical variable, date as a categorical variable for each individual date, and the interaction term between treatment and date. We restricted our measurement period to the days preceding (Nov 9), during (Nov 10-12) and immediately after the heat treatment (Nov 13). We included individual cucumber identity as a random effect to account for repeated measures on the same individuals over time. We also included bucket ID and sea table ID as random effects to account for our paired (two cucumbers per bucket) and blocked (five buckets per sea table) experimental design. We conducted AIC model selection based on the *dredge* function from package MuMIn (CITE) to determine the most parsimonious models (ΔAIC > 2). From top selected models we calculated odds ratios and 95% confidence intervals, and considered effects to be significant if confidence intervals did not overlap zero.

To examine the effect of treatment temperature on the likelihood of a cucumber eviscerating, we used the fitDist function from the {gamlss} R package to select the appropriate distribution and used forward selection to create a logistic regression model with treatment, weight, and pooping status as explanatory variables in the full model, and sea table as a random effect. We ran the model to investigate the significance of the selected explanatory variable(s).

DECLAN

Spawning could only be measured per-bucket instead of per-cucumber because the gametes would not stay only on one side of the bucket housing two cucumbers. We ran an ANOVA test on

To examine the effect of treatment temperature on the likelihood of a cucumber eviscerating, we used the fitDist function from the {gamlss} R package to select the appropriate distribution and used forward selection to create a logistic regression model with treatment, weight, and pooping status as explanatory variables in the full model, and sea table as a random effect. We ran the model to investigate the significance of the selected explanatory variable(s).

We used a boxplot to examine how cucumber activity scores varied between treatments for the days preceding, during and after the heat treatment. We used Kruskal-Wallis non-parametric analysis of variance tests to compare the activity scores of each treatment to assess initial heat shock, initial recovery and long-term recovery. We compared within-treatment activity scores from: 1) the day before the heat treatment and the first day of the heat treatment; 2) the last day of the treatment and the first day after; and 3) the first day after heat treatment and activity scores measured 7 days later.

We then fit a generalized additive model to activity score with a zero-adjusted negative binomial distribution using the *gamlss* function in the GALMSS (CITE) package. This analysis only used data from the days preceding (Nov 9), during (Nov 10-12) and immediately after the heat treatment (Nov 13). We included treatment, date (measured as categorical) and the interaction between treatment and date as fixed effects.