**Statistical Methods**

*Lethal Data*

To examine the effect of treatment on sea cucumber mortality, we used a Kruskal-Wallis test from the R {stats} package to determine if there were significant differences between treatments. We then ran a Dunn test from the R {FSA} package to look at specific treatment comparisons. Finally, we ran a logistic regression model (since our mortality data followed a binomial distribution) to examine the effects of additional parameters on mortality. We used a GAMLSS model with evisceration, defecation status, initial droop and squeeze scores, and initial weight as fixed effects.

*Sublethal Data*

We first examined how cucumber stiffness metrics, indicated by antipredator response and structure maintenance, were correlated 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 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).

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. As such we collected data on which buckets contained stress-spawning cucumbers at some point during the experiment. The data is binomial and non-parametric, so we ran a Kruskal-Wallis test using the {stats} package in R.

R v.4.0.3 was used.

We used Kruskal-Wallis non-parametric analysis of variance tests followed by Dunn tests for multiple comparisons to compare the activity scores between treatments on specific days to assess: pre-experiment differences across treatments (Nov 09), initial heat shock (Nov 10), initial recovery (Nov 13) and long-term recovery (Nov 20). We assessed significance based on alpha = 0.05, and adjusted Dunn test P values for multiple comparisons using Holm’s method. We then fit a generalized additive model with a zero-adjusted negative binomial distribution using the *gamlss* function in the GALMSS (CITE) package. Like the stiffness analysis, we 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. Again, we included individual ID, sea table and bucket ID as random effects. We used backward step-selection to find the most parsimonious model based on GAIC.

*Wasting*

We assessed cucumbers for lesions on the 4th, 5th, 6th and 12th days of the experiment. We classified lesions as either minor or major lesions based on their size and visual appearance (Figure X). We then determined the maximum number of minor and major lesions per cucumber. We assessed whether weight and treatment affected the maximum number of minor lesions using a generalized additive model with a geometric distribution. We included sea table and bucket as random effects.