Factors Controlling Coral Skeletal U/Ca Ratios with Implications for their Use as a Proxy for Past Ocean Conditions

ABSTRACT

An understanding of past environmental conditions provides context for evaluating modern climate changes, but there is a lack of data for periods before the existence of instrument-based environmental records. To reconstruct past conditions, proxies are needed. Recently, the uranium/calcium ratio in fossil coral skeletons has been investigated as a proxy for past ocean conditions. However, existing studies are either lab-based or were developed using small numbers (n~15) of natural samples. This study seeks to explore whether existing calibrations can describe these relationships for a large observational dataset of 835 fossil coral samples. Results suggest that no single variable controls U/Ca, but rather that multiple environmental variables are needed to capture it. Specifically, pH, Ω, TCO_2, alkalinity, and temperature are all significant predictors of U/Ca. Genus also seems to play a key role in these relationships. These results provide further insight into whether the fossil coral uranium proxy can be used to accurately reconstruct past ocean conditions.

INTRODUCTION

Scleractinian corals skeletons are one of the most fruitful paleoclimate archives because they are easily dated, they tend to record environmental properties of the seawater in which they grow, it is possible to make multiple measurements of environmental properties on the same sample due to their relatively large size, and they can capture sub-annual environmental variability.[1, 2] In recent years, the U/Ca ratio of corals has been explored as a proxy (or indirect indicator) for seawater properties including seawater temperature,[3] salinity,[4] pH,[5] and carbonate ion.[6] The ability of U/Ca to capture past seawater salinity, pH, and/or carbonate ion, in particular, would be a valuable addition to the arsenal of coral-based proxies, due to the relative lack of proxies for these environmental variables.

There is reason to expect U/Ca in corals to be sensitive to seawater pH and/or carbonate ion concentrations. Inorganic aragonite precipitation experiments conducted by DeCarlo *et al.* (2015) suggest that U is incorporated into the aragonite mineral in proportion to the ratio of [U]/[CO_3^{2-}] in seawater.[7] These results support the idea that corals, which make aragonite skeletons, may also incorporate uranium in proportion to the concentration of carbonate ion in seawater. A potential problem is the fact that corals do not appear to precipitate their aragonite skeletons directly from seawater, but rather from a semi-enclosed space referred to as the calcifying fluid.[8–10] Because corals are able to manipulate the chemical composition of this calcifying fluid (through active pumping of protons, for example), the calcifying fluid has a carbonate chemistry that is distinct from external seawater.[11–13] As a result, coral U/Ca could be more dependent on other environmental parameters that have strong effects on the coral skeletal growth process.

There is also evidence that coral U/Ca may be dependent on sea surface salinity. First, Swart and Hubbard (1982) showed that coral U/Ca ratios appear to be dependent mainly on the absolute concentration of uranium in seawater.[14] As a result, they postulated that U/Ca in coral could record salinity variations as a result of changes in the absolute concentration of U. Results of Shen and Dunbar (1995), showing that U/Ca in tropical coral skeletons were correlated with the amount of local rainfall at the time the corals grew, generally support this idea.[15]

Many published studies have also examined the potential for relationships between U/Ca and seawater temperature.[3, 16–21] While it does not appear that U/Ca is sensitive to temperature in inorganic aragonite,[7] U/Ca is correlated with temperature in natural samples over seasonal cycles.[22] As a result, U/Ca may be related to seawater temperature through an effect of temperature on the coral skeletal growth process.

To date, most studies calibrating environmental proxies based on coral skeletal U/Ca ratios come from natural samples collected from a single geographic location,[15, 22] or from laboratory-based culture experiments in which one environmental variable is varied while others are held constant.[5] However, if paleoclimatologists hope to apply proxies over a broad geographic range and far into the past (when seawater boundary conditions may have looked quite different), it is important to quantify whether U/Ca depends on a single, or multiple, environmental parameters. It is also important to quantify how precisely environmental variations could be reconstructed from coral U/Ca.

One common approach in proxy development, especially for calibrating proxies in ocean sediment cores, is to conduct a "core-top" calibration. In this approach, measurements of chemical, geological, physical or biological variables are made in the top few centimeters of ocean sediment and are examined for correlations with environmental variables in the water column overlying the sediment core location. Recent studies, for example, have combined core-top measurements of Mg/Ca ratios in foraminifera with oceanographic databases to refine the foraminifera-based Mg/Ca paleothermometer.[23, 24] This study attempts to conduct a core-top calibration of U/Ca's sensitivity to environmental changes by comparing modern and Recent (Holocene and younger) tropical coral U data with environmental data from the World Ocean Atlas (WOA) and the Global Ocean Data Analysis Project (GLODAP).

A wealth of coral U data exists in the literature due to the fact that tropical corals have been measured by U-series dating as a way to generate sample ages for decades (*e.g.* Robinson *et al.* 2014).[25] In order to obtain U-series ages from coral, the abundance of ^{238}U, which makes up >99% of the uranium in the coral skeleton, is determined. As a result, U-series measurements made over the past few decades also produce measurements of coral [U]. Recently, Chutcharavan *et al.* (2018) compiled coral U-series measurements from the literature and analyzed the data to determine whether seawater $^{234}U/^{238}U$ ratios appear to have changed over time.[26] In the present study, the compilation of Chutcharavan *et al.* (2018) is used to investigate whether the [U] of recent corals appear to vary significantly with environmental variables such as temperature, salinity, pH, and/or carbonate ion.

METHODOLOGY

Data Compilation

Coral U-series data compiled in Chutcharavan *et al.* (2018) was used to evaluate the relationship between coral skeleton uranium concentrations and the environmental properties of seawater in which corals grew.[26] The main variables of interest alongside the genus, age, percent calcite, and ^{238}U concentration of the corals include sea surface temperature (SST), salinity (SSS), pH, total alkalinity, dissolved inorganic carbon, and saturation state with respect to aragonite (Ω). A summary of these variables can be found in **Table 1**. All samples for which ^{238}U concentrations were labeled as "not reported" and/or genus was reported as "nd", "unknown", or "unidentified" were excluded from the analysis.

Variable	Minimum	Median	Maximum	Mean
^{238}U (ppm)	0.11	1.16	3.76	1.21
Age (ka)	0	1.27	9.99	2.80
Calcite Abundance (%)	0.10	2.0	23.0	2.91
Salinity (g/kg)	31.13	35.13	36.49	34.82
Temperature (℃)	22.04	26.64	29.49	26.32
pH (total scale)	8.0	8.11	8.15	8.10
Total Alkalinity (μmol/kg)	2176.6	2283.62	2389.54	2277.20
TCO_2 (μmol/kg)	1857.15	1959.35	2047.78	1947.69
Ω	2.87	3.75	4.26	3.70
$\delta^{234}U$ initial (‰)	35.99	144.85	162.97	144.76

Table 1. Data summary of key coral and environmental variables. These values include all non-missing values with age under 10,000 years from the three most abundant genera in the Chutcharavan *et al.* (2018) compilation.

Complementary environmental data was paired with the coral U-series dataset from the World Ocean Atlas (WOA) and the Global Ocean Data Analysis Project (GLODAP). The GLODAP database contains information on Ω, total alkalinity, TCO_2, and pH.[27–29] The WOA database contains climatological information, including both the temperature and salinity of the sea surface.[27,30] Both databases have been updated with recent oceanographic data, taken at gridded points with latitudes, longitudes, and depths associated with them. This feature makes the WOA and GLODAP sets appropriate to work in tandem with one another and with our coral dataset.

To pair coral data from a particular geographic location with environmental data, a distance function was created which measures the distance from each coral sample included in Chutcharavan *et al.* (2018) to the grid locations where each of the environmental variables were measured. Since the GLODAP and WOA data are built on a grid and use interpolation from measured data, coral environmental parameters were estimated by taking the weighted average of the three closest values to each fossil coral sample. The function was coded in C++ for efficiency and run in R using the Rcpp library. The function measured the distance between any two points on earth, that is the coral and each environmental grid location, measured in latitude and longitude, taking the earth's curvature into account. In the case that latitudes or longitudes were not reported in Chutcharavan *et al.* (2018) or in the original scientific papers cited by Chutcharavan *et al.*, the original scientific papers reporting coral U-series data were searched for coral collection sites, and the longitudes and latitudes corresponding to these sites were identified.

Data Cleaning

The coral samples compiled in Chutcharavan *et al.* (2018) were constrained to those of Holocene age (*i.e.* ~10,000 years of age and younger). The Holocene is a time of relatively stable global climate,[31] making it reasonable to pair recent oceanographic environmental data with our coral samples. At ages beyond about 15,000 years, there are fewer coral samples in the compilation for a given time point such that the data are more sparse. Filtering the coral samples to those younger than ~10,000 years results in ^{238}U vs. Age slopes close to zero as seen in **Figure 1** below.

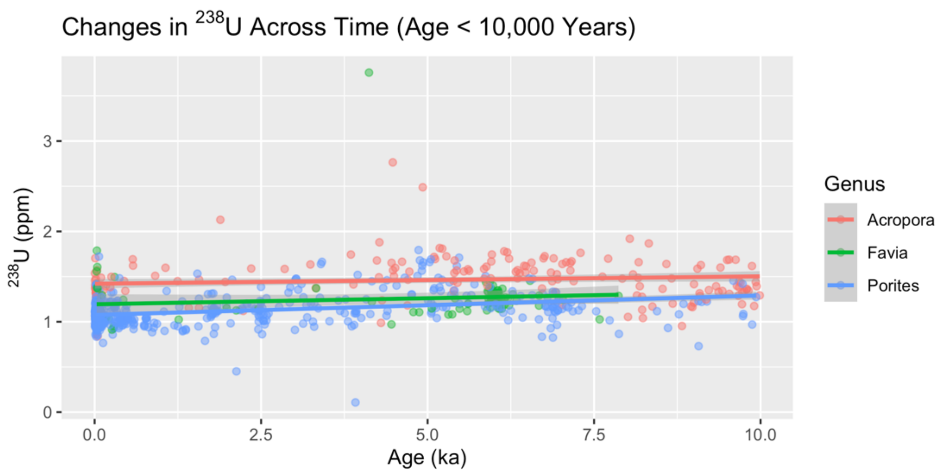

Figure 1. This plot uses coral ^{238}U data as a function of age for samples of the genus Acropora, Favia, and Porites across the entire Chutcharavan *et al.* (2018) dataset, filtered to remove NA's and filtered for samples younger than 10,000 years. Age slopes were close to zero for all genera: Acropora (0.008 (+/- 0.004)), Favia (0.013 (+/- 0.011)), and Porites (0.022* (+/- 0.003)), and when including all three genera (0.03* (+/- 0.003)). *indicates significance at the .05 level. Results include coefficient estimates followed by standard errors.

Variability in relationships between uranium and environmental variables by genus were also examined; only corals with genera that had more than 50 samples were selected. It is notable that 297 samples had no genus listed and were excluded. After selecting the top three genera and filtering for age, 835 samples remained out of the original 1860 (**Table 2**). After data cleaning, 58 sites remained out of the original 78. The original data distribution can be found in **Table A1.**

Genus	Number of Samples	Percent of Total
Porites	543	65%
Acropora	194	23%
Favia	98	12%
All	835	100%

Table 2. After filtering for age and keeping genera with more than 50 samples

Next, the relationship between ^{238}U and % calcite was explored since higher percentages of calcite in the coral skeleton can overprint the original geochemical composition of coral skeletons at the time of growth because uranium incorporation in calcite and aragonite differ.[32] Some labs reported % calcite as a range of values, in which cases the maximum reported value was used in order to avoid underestimating the % calcite of the sample. The data was assessed on variable and genus levels for differences between corals with different amounts of calcite. Calcite was cut into three groups, less than or equal to one percent, greater than one percent, and those with no values listed (≤ 1 = 122 samples, > 1 = 93 samples, N/A = 620 samples). The threshold chosen was one because 1% is the detection limit of many x-ray diffraction measurements that quantify calcite abundance.[26] The directions of the slopes change between different calcite groups for pH and Ω, while there were no significant differences in slopes between groups that reported calcite concentrations for temperature or salinity (**Figure 2, Figure A1**). Some labs in the data that did not record percent calcite may have done so because they were measuring relatively young corals, which are less likely to contain secondary calcite. Since our univariate regression results for temperature, salinity, and pH all did not change significantly when restricting calcite, no calcite restriction was applied in order to obtain a more robust analysis with greater sample sizes.

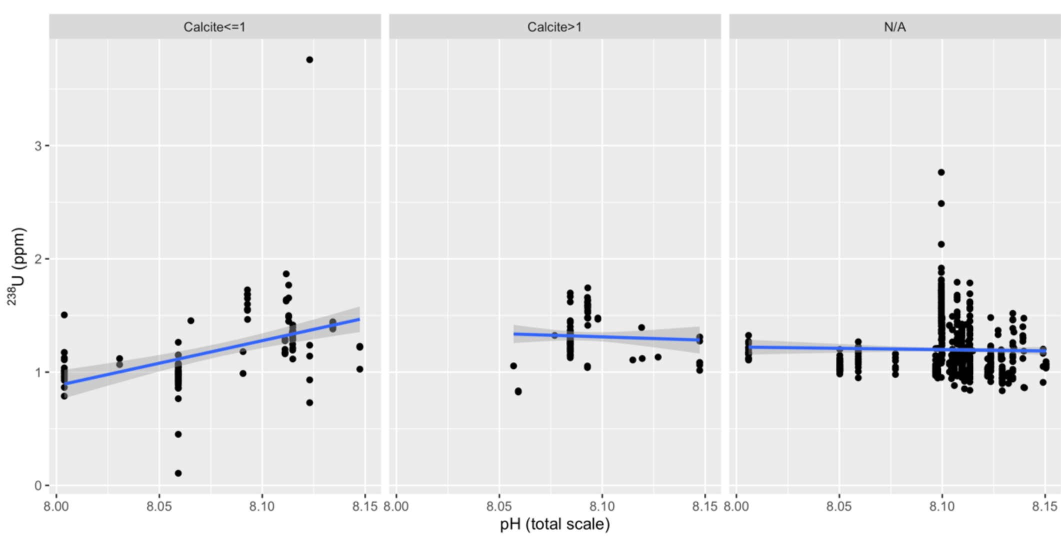

Figure 2. The relationship between ^{238}U and pH for calcite ≤ 1 is significantly different from that of calcite > 1, but not significantly different from samples with no calcite value.

The next data integrity check was to examine ^{238}U values at the lab level for outliers. After examining labs with more extreme means and taking into account environmental conditions and sample size, no exclusions attributable to outlying lab data were necessary (**Figure 3**).

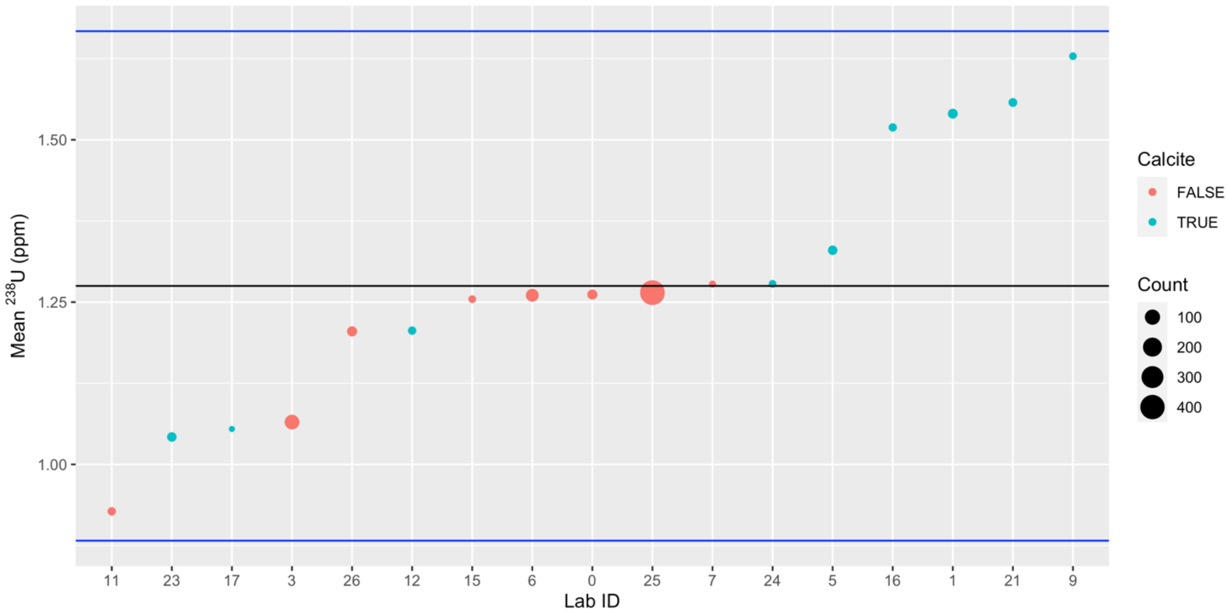

Figure 3. Each lab's mean ^{238}U was calculated for all samples, and points sized by number of samples. Points are colored by whether or not the lab reported calcite values for their sample. The black horizontal line is the global mean, and the blue lines are +- two standard errors from the lab aggregated mean.

An additional analysis conducted to validate the data was to see if any locations had unusual or strongly correlated environmental data values. Since environmental data is resolved at the location level, there are only 32 locations and thus 32 values for each of the environmental variables (temperature, salinity, pH, Ω, alkalinity, TCO_2). There were a few points of interest that had particularly low salinity values compared to their pH and temperature values, as seen in **Figure 4**.

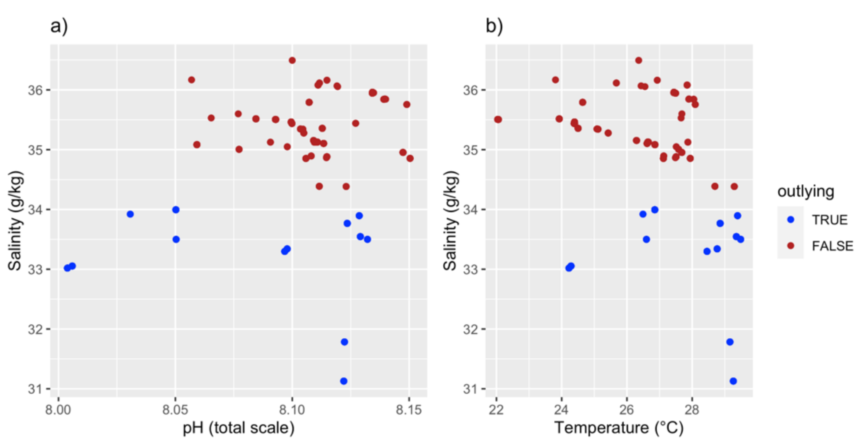

Figure 4. Low salinity points plotted against pH and temperature.

Upon further inspection, these points were valued as expected. They came from coastal regions in East Asia which are known to have lower salinity values due to the proximity of freshwater rivers, visualized in **Figure 5**.

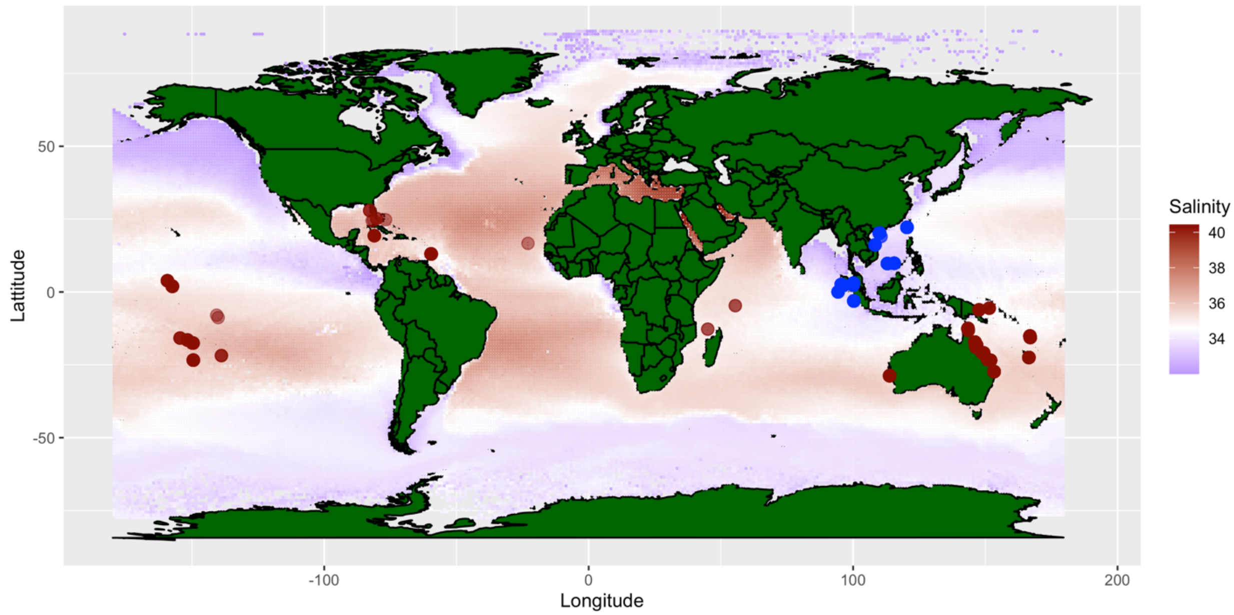

Figure 5. Global map with ocean salinity colored from blue (low) to red (high), with corresponding outlying points marked by location.

Statistical Methods
Univariate linear regression was used to test if ^{238}U can be used as a proxy for temperature, salinity, and pH separately, as has been done in previous studies.[3, 6, 16–22, 33, 34] Additionally, multiple linear regression was used in order to account for additional variables which cannot be controlled for in an observational setting. In order to perform multiple linear regression, predictors need to be chosen, a process known as feature selection. Regularized regression (*e.g.* lasso and ridge regression) has been widely used in recent years due to its demonstrated prediction prowess on data with many predictors of unknown relationships. Regularized regression is suited to this task because the underlying relationships are expected to be linear, not categorical or tree-like.

To determine which type of regularized regression to use, the data was divided into 90% training data, and 10% test data for validation purposes. Lasso, ridge, and elastic net regression were compared, using 10-fold cross-validation to select the minimum and one-standard error penalty parameters. Then, the prediction of ^{238}U in the test data was compared by root mean squared error (RMSE) across the three modeling techniques. Lasso was chosen because it had the lowest root mean squared error when predicting on the validation data.

Bootstrapping and bagging were employed to ensure stable estimates from variable selection. Bootstrap aggregation, or bagging, was performed for the lasso modeling. The data was bootstrapped 100 times, a lasso model was fit to the bootstrap sample, and relevant information was recorded.

Best subset selection was implemented for comparison with the bootstrapped lasso approach. This involved fitting every possible combination of linear models with a given set of predictors, of which there are 2^p possible models, where p is the number of predictors in the data. These models can then be compared with a selection of characteristics, such as adjusted R-squared, or information criterion like AIC (Akaike information criterion) and BIC (Bayesian information criterion). BIC was selected since it more heavily penalizes additional predictors than AIC, and it was valuable to achieve a parsimonious model to compare against the results of bootstrap lasso.

Finally, principal component analysis (PCA) was performed on the data to help understand the relationships between the predictors. PCA allows the collapse of high dimensional data into a few independent components. Singular value decomposition is applied to the input data and results in a linear transformation of the data such that the first new predictor describes most of the variability in the data, and each subsequent predictor describes less. The original predictors can also be visualized in "PCA-space" to give some understanding of how they relate to one another.

RESULTS

Comparison to Previous Studies

Temperature

Linear regression was used to test if there was a significant association between explanatory variables and ^{238}U, when applying an age restriction of 10,000 years, and no restriction based on percent calcite. The first test was between temperature and ^{238}U (**Figure A2**). When including all three species of interest, there was significant evidence of a negative association between temperature and ^{238}U ($t = -15.40$, p-value < 0.001). When looking at each individual genus, a highly significant negative association between temperature and ^{238}U was found for Porites ($t = -10.63$, p-value < 0.001) and for Acropora ($t = -7.67$, p-value < 0.001). There was an insignificant, but surprisingly positive association between temperature and ^{38}U for Favia ($t = 1.62$, p-value $= 0.109$). Negative relationships between ^{238}U and temperature are consistent with results from previous studies on Porites coral (**Table 3**, **Figure 6**). Comparisons between the results in this study with and without filtering for age and calcite can be found in **Tables A2** and **A3**.

Our Results			Previous Lab Studies			
Intercept	**Temperature**	**Genus**	**Intercept**	**Temperature**	**Genus**	**Source**
2.283* (+/- 0.110)	-0.043* (+/- 0.004)	Porites	1.928	-0.033*	Porites	Ourbak et al. (2006)
2.965* (+/- 0.197)	-0.059* (+/- 0.008)	Acropora	2.232	-0.0456*	Porites	Min et al. (1995)
0.330 (+/- 0.571)	0.037 (+/- 0.023)	Favia	1.963	-0.032*	Porites	Quinn and Sampson (2002)
2.873* (+/- 0.108)	-0.063* (+/- 0.004)	All Genera	2.106	-0.0367*	Porites	Correge et al. (2000)
			1.488	-0.0212*	Porites	Armid et al. (2011)
			1.957	-0.029*	Porites	Wei et al. (2000)
			2.057	-0.034*	Porites	Felis et. al. (2009)
			2.26	-0.044*	Porites	Fallon et al. (1999)
			2.24	-0.046*	Porites	Sinclair et al. (1998)

Table 3. Our results, which showed a negative association between temperature and ^{238}U, compared to previous lab studies. *indicates significance at the .05 level. Our results include coefficient estimates followed by standard errors.

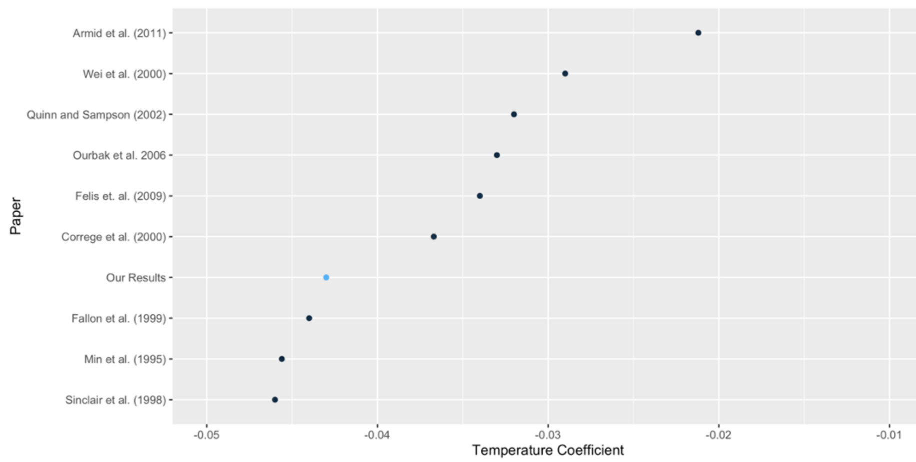

Figure 6. Coefficients from our result using only Porites compared to other reported Porites slopes using univariate OLS regression.

The next set of models examined the relationship between temperature and ^{238}U after accounting for salinity, since salinity is more variable in this large-scale observational study than in lab studies. The model with all three genera revealed that a significant negative relationship between temperature and ^{238}U remained after adjusting for salinity ($t = -10.60$, p-value < 0.001). When breaking the analysis down by genus, this negative relationship was significant for Porites ($t = -7.77$, p-value < 0.001), and for Acropora ($t = -6.85$, p-value < 0.001), but not for Favia ($t = -1.50$, p-value $= 0.14$), whose relationship was insignificantly positive before accounting for salinity. After accounting for salinity, the relationships between temperature and ^{238}U were negative across all genera. Compared to Ourbak *et al.* (2006), both the Porites model and larger model including all three genera in this study followed a similar pattern, revealing a significant negative relationship between temperature and ^{238}U after accounting for salinity.

After accounting for temperature, the model with all three genera showed a positive effect of salinity on ^{238}U ($t = 5.91$, p-value < 0.001). However, when breaking the analysis down by genus, this positive effect of salinity after adjusting for temperature was only significant for Porites ($t = 4.65$, p-value < 0.001), but not for Acropora ($t = 0.96$, p-value $= 0.338$). There was a significant relationship between salinity and ^{238}U after accounting for temperature for Favia ($t = -2.94$, p-value $= 0.004$), although this relationship was negative. Results for salinity in Porites revealed a significant positive relationship between salinity and ^{238}U after adjusting for temperature, which is also consistent with Ourbak *et al.* (2006) (**Table 4**).

Ourbak et al. (2006)			
Intercept	**Temperature**	**Salinity**	**Genus**
0.162 (+/- 0.019)	-0.022* (+/- 0.003)	0.162* (+/- 0.019)	Porites
Our Results			
Intercept	**Temperature**	**Salinity**	**Genus**
0.958* (+/- 0.305)	-0.034* (+/- 0.004)	0.031* (+/- 0.007)	Porites
1.239 (+/- 1.808)	-0.057* (+/- 0.008)	0.047 (+/- 0.049)	Acropora
25.180* (+/- 8.477)	-0.059 (+/- 0.040)	-0.635* (+/- 0.216)	Favia
0.659 (+/- 0.389)	-0.049* (+/- 0.005)	0.053* (+/- 0.09)	All Genera

Table 4. Our results for temperature and salinity compared to previous lab studies.
*indicates significance at the 0.05 level. Our results include coefficient estimates followed by standard errors.

pH

Linear regression was also used to test whether pH was significantly associated with ^{238}U (**Figure A2**). For the model including all three genera, there was evidence that pH was significantly positively associated with ^{238}U (t = 2.74, p-value = 0.006). When looking at each individual genus, pH was significantly related to ^{238}U only for Porites (t = 2.38, p-value = 0.017) and Acropora (t = -5.94, p-value < 0.001), although this relationship was positive for Porites and negative for Acropora. A positive yet insignificant relationship was found between pH and ^{238}U for Favia (t = 0.15, p-value = 0.88). The negative relationship between pH and ^{238}U in the Acropora model was similar to the results of Inoue *et al.* (2011) which also modeled only Acropora. However, the magnitude of our results was about 25 times larger than was seen in Inoue *et al.* for the intercept, and about 43 times larger for the slope (**Table 5**). After accounting for temperature or temperature and salinity in the Acropora model, the intercept was large but insignificant, and the model revealed no significant relationship between ^{238}U and pH in Acropora (**Table 6**).

Inoue et al. (2011)		
Intercept	**pH**	**Genus**
2.96	-0.21*	Acropora
Our Results		
Intercept	**pH**	**Genus**
-2.740 (+/- 1.620)	0.476* (+/- 0.200)	Porites
74.496* (+/- 12.298)	-9.011* (+/- 1.517)	Acropora
-1.257 (+/- 16.744)	0.310 (+/- 2.068)	Favia
-5.113* (+/- 2.307)	0.781* (+/- 0.285)	All Genera

Table 5. Results for pH (filtering for age and stratified by genus) compared to previous lab studies. *indicates significance at the 0.05 level. Our results include coefficient estimates followed by standard errors.

pH Acropora models			
Intercept	**pH**	**Temperature**	**Salinity**
74.496* (+/- 12.298)	-9.011* (+/- 1.517)		
14.897 (+/- 17.663)	-1.491 (+/- 2.207)	-0.053* (+/- 0.012)	
17.513 (+/- 17.796)	-2.082 (+/- 2.265)	-0.048* (+/- 0.013)	0.057 (+/- 0.050)

Table 6. Comparing univariate pH model to multivariate models for Acropora. *indicates significance at 0.05 level. Our results include coefficient estimates followed by standard errors.

Further Analysis - Variable Selection

In this observational study, ^{238}U is not adequately explained by one or two variables as described in previous studies. Multiple regression modeling can be utilized in this context as an attempt to replicate how certain conditions can be carefully controlled in lab studies (**Table 7**). Four multiple regression models were built, one for each genus subset of corals, and one overall, controlling for all six variables of interest simultaneously. These multiple regression models confirmed some relationships the univariate models in this study have previously shown; for example, pH has a positive relationship with ^{238}U among all genera, modeled individually and overall, when controlling for temperature, Ω, total alkalinity, salinity and total CO_2.

Interestingly, this positive relationship is significant in Acropora when controlling for all other parameters, contrary to the inverse relationship found between pH and ^{238}U in the previous results (**Tables 6** and **7**). The relationship between temperature and ^{238}U is negative overall, after accounting for the other five parameters. The relationship between temperature and ^{238}U is significant in all three genera modeled individually when controlling for the other five parameters, but differs in direction, where Porites and Favia have a negative relationship, and Acropora has a positive relationship. This is again different from the previous result where only salinity was controlled for, in which all genera showed a negative relationship between ^{238}U and temperature. Only pH, Ω, and total alkalinity are significantly related to ^{238}U in Favia, whereas all terms except salinity are significant in the model that includes all three genera.

	Porites	**Acropora**	**Favia**	**All Genera**
Intercept	-41.062* (+/-12.598)	-180.802* (+/-83.972)	-516.199* (+/-212.986)	-49.335* (+/- 16.366)
pH	5.130 *(+/-1.493)	21.864* (+/-10.028)	62.105* (+/-24.208)	6.173* (+/- 1.936)
Temperature	-0.043* (+/-0.008)	-0.031 (+/-0.029)	0.678* (+/-0.170)	-0.054* (+/- 0.008)
Ω	-0.661* (+/-0.189)	-3.341* (+/-0.679)	-20.590* (+/-4.596)	-1.014* (+/- 0.236)
Alkalinity	0.004* (+/-0.001)	0.015* (+/-0.008)	0.140* (+/-0.020)	0.008* (+/- 0.002)
Salinity	-0.032* (+/-0.013)	0.109 (+/-0.074)	-0.891 (+/-0.461)	-0.031 (+/- 0.017)
TCO_2	-0.002 (+/-0.001)	-0.010 (+/-0.010)	-0.111* (+/-0.017)	-0.006* (+/- 0.002)

Table 7. Multiple Linear Regression with all variables for all three major genera.
*indicates significance at the 0.05 level. Our results include coefficient estimates followed by standard errors.

While it is beneficial to see which environmental parameters are most significantly related to ^{238}U when including all variables of interest in multiple linear regression, feature selection can be used to further identify the subset of predictors that are most significantly related to the response, ^{238}U. First, principal component analysis was performed to identify how predictors are related to one another (**Figure 7**, **Figure A3**). The first PCA component captures 52.4% of the variability in the explanatory variables, and the second captures 38.3% (**Figure A4**). This means that the set of explanatory variables can be projected into two dimensions while capturing 90.7% of the variability in the data. One notable piece of information derived from PCA is variable loading, or how strongly predictors align with one PCA dimension. It is clear that pH is very highly loaded on the second dimension because it points nearly parallel to that axis. Salinity and total alkalinity are highly loaded on dimension one, though to a lesser extent than pH is loaded on dimension two. Finally, temperature and TCO_2 measure a very similar parameter, but are inversely proportional.

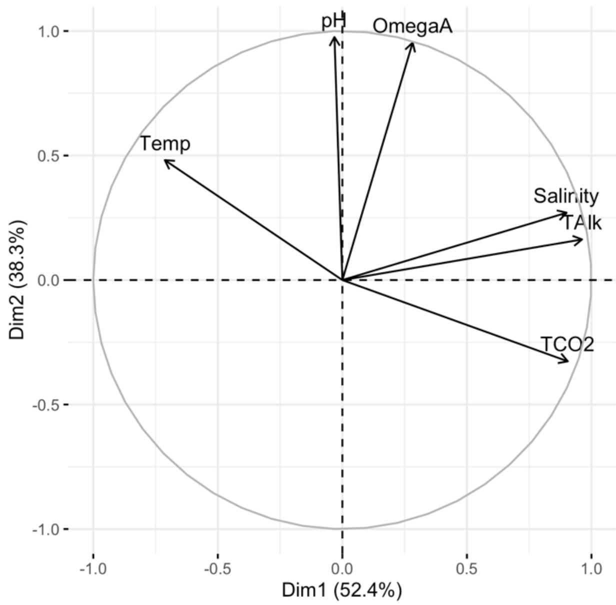

Figure 7. Visualization of predictors projected onto the first two PCA dimensions.

To perform feature selection, lasso and bagging of lasso were employed, the latter of which accounted for instabilities in lasso estimates. The variables of interest in these analyses were temperature, salinity, pH, Ω, total alkalinity, and total CO_2. Because lasso identifies the penalty parameter using cross-validation, the chosen variables and corresponding slopes change slightly each time the process is executed. To account for this variability, bagging was performed on the lasso models. The data was bootstrapped 100 times. For each of those times, lasso performed variable selection, and minimum penalty parameters and slopes of the shrunken variables were recorded (**Figure 8**).

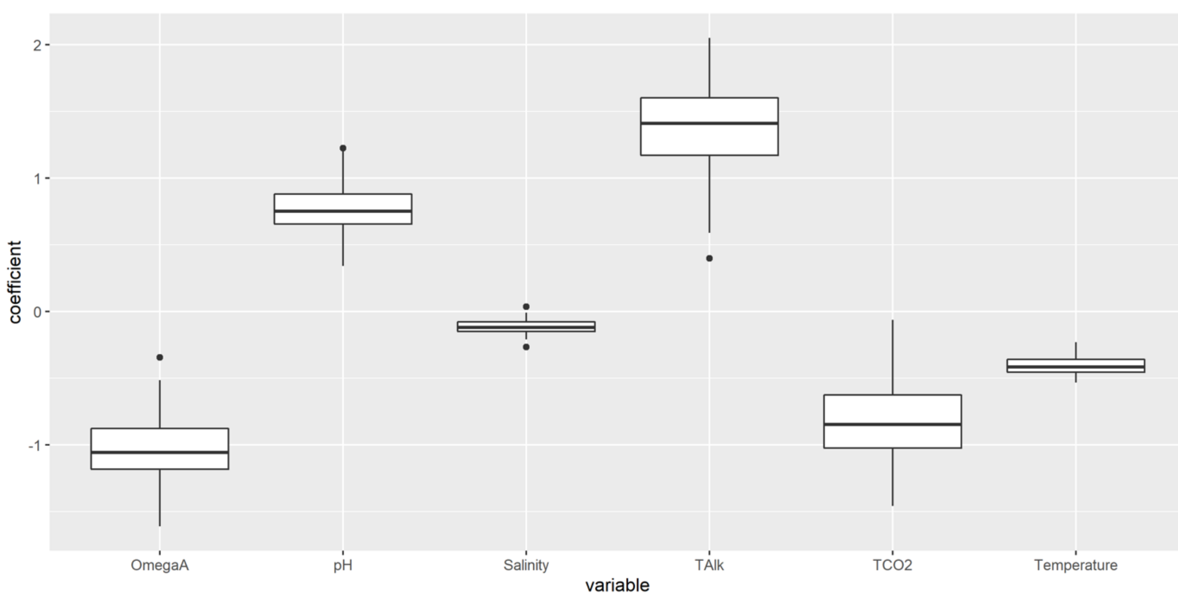

Figure 8. Box plots of slopes of standardized variables resulting from minimum penalty parameter lasso run on 100 bootstraps of the data. Lasso sets unimportant variables to zero, thus plots that are centered or overlapping with zero are considered unimportant, while those that do not overlap zero are selected.

It can be seen that when using the minimum error penalty parameter, Ω, pH, TCO_2, temperature and alkalinity were frequently non-zero (that is, selected by the model). Best subset selection identified the same subset of predictors except alkalinity when using all possible combinations of predictors and BIC as the criterion. AIC additionally included salinity. The second best model using both AIC and BIC as the criteria exactly matched our results using bootstrap lasso. Thus, results indicate that Ω, pH, TCO_2 temperature, and alkalinity all control ^{238}U to some extent, and a final multiple linear regression was built using all five as predictors for all genera to quantify these relationships. This final model reveals that ^{238}U has inverse relationships with temperature, Ω, and TCO_2 and direct relationships with pH and alkalinity (**Table 8**). The inverse relationship between temperature and ^{238}U is consistent with previous calibrations. While the relationship between pH and ^{238}U is not consistent with past studies, this is likely due to measuring the effect of pH while controlling Ω, which is very difficult to measure in practice as these variables are closely coupled. These relationships also likely differ by genus.

	All Genera
Intercept	-58.261* (+/- 15.627)
pH	7.208* (+/- 1.852)
Temperature	-0.048* (+/- 0.008)
Ω	-1.094* (+/- 0.232)
TCO_2	-0.004* (+/- 0.002)
Alkalinity	0.007* (+/- 0.002)

Table 8. Multiple linear regression with variables selected from lasso and best subset selection. *indicates significance at the 0.05 level. Our results include coefficient estimates followed by standard errors.

DISCUSSION

This study reveals that using a large observational data set of coral fossils yields results consistent with several existing, lab-based studies measuring temperature and salinity with smaller sample sizes. Specifically, this study confirms previous lab calibrations that describe the univariate relationship between ocean temperature and ^{238}U in fossil coral. After accounting for salinity, the negative relationship remains between temperature and ^{238}U. Additionally, after adjusting for temperature, a significant positive relationship was found between salinity and ^{238}U. Both relationships are consistent with the results of Ourbak *et al.* (2006).[4] Interestingly, the relationship between pH and ^{238}U differed from past lab studies. Specifically, an overall positive relationship between pH and ^{238}U was found, inconsistent with results from lab studies such as Inoue *et al.* (2011).[5] This discrepancy was found in all of the univariate models except for one which only modeled Acropora. However, after adjusting for temperature, there was essentially no relationship between pH and ^{238}U in the Acropora model. This may be reconciled in two ways. First, univariate relationships are generally insufficient to describe this observational data. Secondly, existing lab studies are not able to decouple pH and Ω. This analysis reveals a negative relationship between Ω and ^{238}U, consistent with past studies. Indeed, previous univariate model results showing pH as a good predictor of ^{238}U may be related to pH's high loading on PCA dimension 2. The observation that pH corresponds nearly exactly with a PCA dimension shows that it likely captures some information from other parameters that are also highly loaded on PCA dimension 2 like Ω and temperature. Thus, slopes reported for pH from previous experiments may indicate control by either Ω and temperature and not by pH itself.

Based on this work, univariate relationships are insufficient to fully describe the complex interactions of corals and seawater conditions. This result suggests that it would likely be challenging to apply a simple proxy for inferring past oceanic conditions as represented by coral ^{238}U, since many parameters control ^{238}U. However, if there is evidence that other variables remain constant in some regions or time periods, it could be possible to predict other variables from ^{238}U. Specifically, since the relationship between temperature and ^{238}U remained even after accounting for salinity, it is possible that ^{238}U could be used as a proxy for measuring temperature. The relationship between pH and ^{238}U does not appear as simple.

A major difference with this observational data compilation and lab studies is that lab studies often fix all experimental parameters besides the one or two they are studying. Furthermore, a study to determine which combination of many possible parameters best describe ^{238}U in corals has not been done. However, the difficulty is that in observational ocean data, and in many lab experiments, environmental parameters may covary. With careful variable selection, this difficulty can be somewhat addressed by statistically approximating the control of parameters using multiple linear regression, although when two parameters are highly correlated it is difficult to evaluate the effect of changes in one while holding the other constant. When including all parameters of interest in variable selection through bootstrap lasso and best subset selection, pH, Ω, TCO_2, alkalinity, and temperature are all significant predictors of ^{238}U. Genus also seems to play a key role in these relationships.

It would be valuable to further test these results in additional laboratory experiments due to some of the limitations of the dataset. One limitation was the need to filter by genus (rather than by species) to include large enough sample sizes. Hence, there could be species-specific effects that are not captured. Additionally, climate was assumed to be relatively constant over the last 10,000 years, but previous work has highlighted regional changes in climate over this time period that are unaccounted for in this examination.[35] Furthermore, many of the environmental variables of interest are highly correlated, which can make regression results difficult to interpret. Finally, the range in the environmental variables is relatively small (*e.g.* pH values from 8.00 to 8.15), especially compared to lab calibrations. Though significant conclusions can still be drawn, values over larger data ranges, as can be obtained through further lab experiments, would help to more fully understand the relationships identified in this study.

CONCLUSIONS

The abundance of uranium in corals has been recently explored as a potential proxy for seawater temperature, salinity, pH, and carbonate ion concentrations. However, most previous studies on uranium's response to environmental change either come from lab calibrations, in which a single environmental variable is varied while others are held constant, or field studies in which uranium's sensitivity to environmental change is examined at one or two geographic locations. In this study, a compilation of U-series measurements in tropical corals from a range of geographic locations was used in combination with environmental variables from two oceanographic databases to quantify relationships between coral ^{238}U and seawater temperature, salinity, aragonite saturation state, and pH. Univariate linear regressions and multiple linear regressions were used to compare relationships between uranium and environmental parameters. Results of these analyses indicate that uranium is dependent on multiple environmental parameters and that previously developed univariate regressions may be insufficient to characterize the full range of variables that influence coral ^{238}U. In addition, relationships between ^{238}U and environmental variables vary by genus. Further laboratory experiments, in which larger ranges of environmental variability can be explored, may prove useful in further testing the multivariate relationships found here and for identifying the physical, chemical, and biological mechanisms driving the dependences of coral uranium abundance on environmental change.

REFERENCES

1. Druffel, E. (1997) Geochemistry of corals: Proxies of past ocean chemistry, ocean circulation, and climate. *Proc. Natl. Acad. Sci.* 94, 8354–8361.

2. Edwards, R.L., Gallup, C.D., Cheng, H. (2003) Uranium-series Dating of Marine and Lacustrine Carbonates. *Reviews in Mineralogy and Geochemistry* 52, 363–405.

3. Correge, T., Delcroix, T., Recy, J., Beck, W., Cabioch, G., Le Cornec, F. (2000), Evidence for stronger El Nino-Southern Oscillation (ENSO) events in a mid-Holocene massive coral, *Paleoceanography* 15, 465–470.

4. Ourbak, T., Correge, T., Malaize, B., Le Cornec, F., Charlier, K., and Peypouquet, J.-P. (2006) A high-resolution investigation of temperature, salinity, and upwelling activity proxies in corals. *Geochem. Geophys. Geosyst.* 7, Q03013.
5. Inoue, M., Suwa, R., Suzuki, A., Sakai, K., and Kawahata, H., (2011). Effects of seawater pH on growth and skeletal U/Ca ratios of *Acropora digitifera* coral polyps. *Geophysical Research Letters* 38, 12801–12804.
6. Anagnostou, E., Sherrell, R. M., Gagnon, A. C., LaVigne, M., Field, M. P., and McDonough, W. F., (2011). Seawater nutrient and carbonate ion concentrations recorded as P/Ca, Ba/Ca, and U/Ca in the deep-sea coral *Desmophyllum dianthus*. *Geochimica et Cosmochimica Acta* 75, 2529–2543.
7. DeCarlo, T.M., Gaetani, G.A., Holcomb, M., Cohen, A.L., (2015). Experimental determination of factors controlling U/Ca of aragonite precipitated from seawater: Implications for interpreting coral skeleton. *Geochimica et Cosmochimica Acta* 162, 151–165.
8. McConnaughey, T. (1989) ^{13}C and ^{18}O isotopic disequilibrium in biological carbonates: II. *In vitro* simulation of kinetic isotope effects. *Geochimica et Cosmochimica Acta* 53, 163–171.
9. Adkins J. F., Boyle E. A., Curry W. B. and Lutringer A. (2003) Stable isotopes in deep-sea corals and a new mechanism for "vital effects." *Geochim. Cosmochim. Acta* 67, 1129–1143.
10. Tambutté, E., Tambutté, S., Segonds, N., Zoccola, D., Venn, A., Erez, J. and Allemand, D. (2012) Calcein labelling and electrophysiology: insights on coral tissue permeability and calcification. *Proc. R. Soc. B Biol. Sci.* 279, 19–27.
11. Comeau, S., Tambutte, E., Carpenter, R., Edmunds, P., Evensen, N, Allemand, C., Ferrier-Pages, C., Tambutte, S., Venn A. (2017) Coral calcifying fluid pH is modulated by seawater carbonate chemistry not solely seawater pH. *Proc. R. Soc. B Biol. Sci.* 284, 20161669.
12. Al-Horani, F.A., Al-Moghrabi, S.M., de Beer, D. (2003) The mechanism of calcification and its relation to photosynthesis and respiration in the scleractinian coral *Galaxea fascicularis*. 142, 419-426.
13. Zoccola D., Tambutté E., Kulhanek E., Puverel S., Scimeca J.-C., Allemand D. and Tambutté S. (2004) Molecular cloning and localization of a PMCA P-type calcium ATPase from the coral Stylophora pistillata. *Biochim. Biophys. Acta - Biomembr.* 1663, 117–126.
14. Swart, P. K., and Hubbard, J. A. E. B., (1982). Uranium in Scleractinian Coral Skeletons. *Coral Reefs* 1, 13–19.
15. Shen, G. T., and Dunbar, R. B., (1995). Environmental controls on uranium in reef corals. *Geochimica et Cosmochimica Acta* 59, 2009–2024.
16. Quinn, T.M. and Sampson, D.E. (2002) A multiproxy approach to reconstructing sea surface conditions using coral skeleton geochemistry, *Paleoceanography and Paleoclimatology* 17, doi:10.1029/2000PA000528.
17. Armid, A., Asami, R., Fahmiati, T., Sheikh, M.A., Fujimura, H., Higuchi, T., Tair, E., Shinjo, R., Oomori, T. (2011) Seawater temperature proxies based on DSr, DMg, and DU from culture experiments using the branching coral *Porites cylindrica*, *Geochem. Cosmochim. Acta* 75, 4273–4285.
18. Wei, G., Sun, M., Li, X., Nie, B. (2000) Mg/Ca, Sr/Ca, and U/Ca ratios of a porites coral from Sanya Bay, Hainan Island, South China Sea and their relationships to sea surface temperature, *Paleogeog., Paleoclim., Paleoecol.* 162, 59–74.
19. Felis, T., Suzuki, A., Kuhnert, H., Dima, M., Lohmann, G., Kawahata, H. (2009) Subtropical coral reveals abrupt early-twentieth-century freshening in the western North Pacific Ocean *Geology* 37, 527–530.
20. Fallon, S.J., McCulloch, M.T., van Woesik, R., Sinclair, D.J. (1999) Corals at their latitudinal limits: laser ablation trace element systematics in *Porites* from Shirigai Bay, Japan. *Earth Planet. Sci. Lett.* 172, 221–238.
21. Sinclair, D.J., Kinsley, L.P.J., McCulloch, M.T. (1998) High resolution analysis of trace elements in corals by laser ablation ICP-MS, *Geochimica et Cosmochimica Acta* 62, 1889-1901.
22. Min, G. R., Edwards, R. L., Taylor, F. W., Recy, J., Gallup, C. D., and Beck, J. W., (1995). Annual cycles of U/Ca in coral skeletons and U/Ca thermometry. *Geochimica et Cosmochimica Acta* 59, 2025–2042.
23. Saenger, C.P., Evans, M.N. (2019) Calibration and Validation of Environmental Controls on Planktic Foraminifera Mg/Ca Using Global Core-Top Data. *Paleoceanography and Paleoclimatology* 34, 1249–1270.
24. Tierney, J.E., Malevich, S.B., Gray, W., Vetter, L., Thirumalai K. (2019) Bayesian Calibration of the Mg/Ca Paleothermometer in Planktic Foraminifera, *Paleoceanography and Paleoclimatology* 34, 2005–2030.

25. Robinson, L.F., Adkins, J.F., Frank, N., Gagnon, A.C., Prouty, N.G., Roark, E.B., van de Flierdt, T. (2014) The geochemistry of deep-sea coral skeletons: A review of vital effects and applications for paleoceanography. *Deep Sea Research Part Ii: Topical Studies in Oceanography* 99, 184-198.
26. Chutcharavan, P. M., Dutton, A., Ellwood, M. J., (2018). Seawater $^{234}U/^{238}U$ recorded by modern and fossil corals. *Geochimica et Cosmochimica Acta* 224, 1–17.
27. Locarnini, R. A., A. V. Mishonov, O. K. Baranova, T. P. Boyer, M. M. Zweng, H. E. Garcia, J. R. Reagan, D. Seidov, K. Weathers, C. R. Paver, and I. Smolyar, (2018). *World Ocean Atlas 2018, Volume 1: Temperature.* A. Mishonov Technical Ed.; NOAA Atlas NESDIS 81, 52 pp.
28. Lauvset, S.K., Key, R.M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterstrom, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F.F., Watelet, S. (2016) A new global interior ocean mapped climatology: the 1 x 1 GLODAP version 2, *Earth Syst. Sci. Data* 8, 325–340.
29. Olsen, A., Key, R.M., van Heuven, S., Lauvset, S.K., Velo, A., Lin, Z., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterstrom, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F.F., Suzuki, T. (2016). The Global Ocean Data Analysis Project version 2 (GLODAPv2) - an internally consistent data product for the world ocean *Earth Syst. Sci. Data* 8, 297–323.
30. Zweng, M. M., J. R. Reagan, D. Seidov, T. P. Boyer, R. A. Locarnini, H. E. Garcia, A. V. Mishonov, O. K. Baranova, K. Weathers, C. R. Paver, and I. Smolyar, (2018) *World Ocean Atlas 2018, Volume 2: Salinity.* A. Mishonov Technical Ed.; NOAA Atlas NESDIS 82, 50 pp.
31. GRIP Members; Anklin, M., Barnola, J.M., Beer, J., Blunier, T., Chappellaz, J., Clausen, H.B., Dahl-Jensen, D., Dansgaard, W., de Angelis, M., Delmas, R. J. (1993) Climate instability during the last interglacial period recorded in the GRIP ice core. *Nature* 364, 203–207.
32. Gothmann, A.M., Higgins, J.A., Adkins, J.F., Broecker, W., Farley, K.A., McKeon, R., Stolarski, J., Planavsky, N., Wang, X., Bender, M.L. (2019) A Cenozoic record of seawater uranium in fossil corals. *Geochim. et Cosmochim. Acta* 250, 173–190.
33. Armid, A., Takaesu, Y., Fahmiati, T., Yoshida, S., Hanashiro, R., Fujimura, H., Higuchi, T., Taira, E., and Oomori, T., (2008). U/Ca as a possible proxy of carbonate system in coral reef *in Proceedings of the 11th International Coral Reef Symposium*, 92–96.
34. Inoue, M., Nakamura, T., Tanaka, Y., Suzuki, A., Yokoyama, Y., Kawahata, H., Sakai, K., Gussone, N. (2018) A simple role of coral-algal symbiosis in coral calcification based on multiple geochemical tracers. *Geochim. Cosmochim. Acta* 235, 76–88.
35. Mayewski, P.A., Rohling, E.E., Stager, J.C., Karlen, W., Maasch, K.A., Meeker, L.D., Meyerson, E.A., Gasse, F., van Kreveld, S., Holmgren, K., Lee-Thorp, J., Rosqvist, G., Rack, F., Staubwasser, M., Schneider, R.R., Steig, E.J. (2004) Holocene climate variability. *Quaternary Research* 62, 243–255.

APPENDIX

Species	Number of Samples	Percent of Total
Porites	713	38.3%
Acropora	459	24.7%
Favia	137	7.4%
Montastraea	80	4.3%
Other	174	9.4%
N/A	297	16.0%

Table A1. Original data prior to data cleaning.

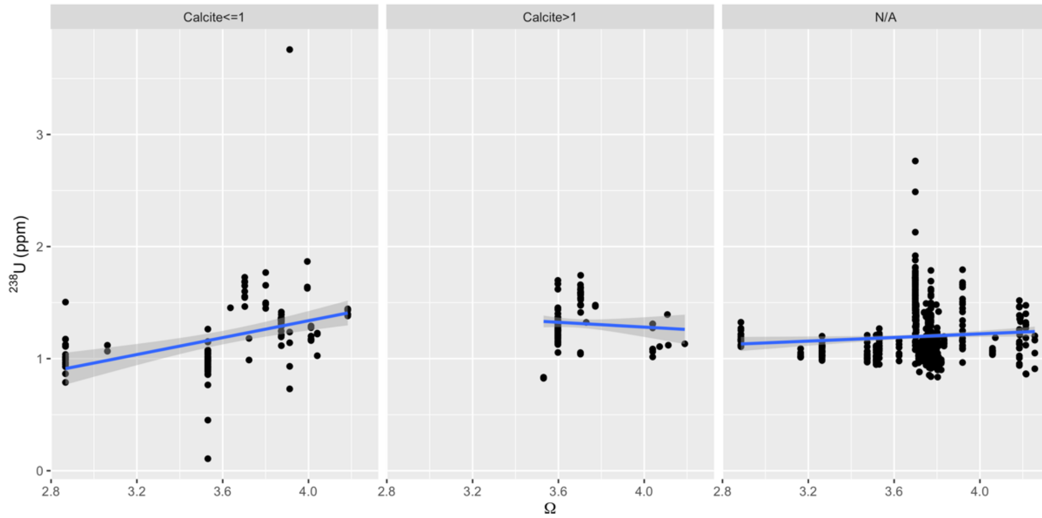

Figure A1. The relationship between ^{238}U and Ω for calcite ≤ 1 is significantly different from that of calcite > 1, but not significantly different from samples with no calcite value.

Primary Results				No Age Restriction			
Intercept	**Temperature**	**Genus**	**Count**	**Intercept**	**Temperature**	**Genus**	**Count**
2.283* (+/- 0.110)	-0.043* (+/- 0.004)	Porites	543	2.070* (+/- 0.113)	-0.034* (+/- 0.004)	Porites	713
2.965* (+/- 0.197)	-0.059* (+/- 0.008)	Acropora	194	2.733* (+/- 0.128)	-0.050* (+/- 0.005)	Acropora	459
0.330 (+/- 0.571)	0.037 (+/- 0.023)	Favia	98	1.662* (+/- 0.333)	-0.017 (+/- 0.013)	Favia	137
2.873* (+/- 0.108)	-0.063* (+/- 0.004)	All Genera	835	2.471* (+/- 0.099)	-0.046* (+/- 0.004)	All Genera	1309

Table A2. Comparison of our results for temperature with and without age restriction of 10,000 years.
*indicates significance at the .05 level

Primary Results				No Age Restriction			
Intercept	**Temperature**	**Genus**	**Count**	**Intercept**	**Temperature**	**Genus**	**Count**
2.283* (+/- 0.110)	-0.043* (+/- 0.004)	Porites	543	1.313* (+/- 0.475)	-0.012 (+/- 0.084)	Porites	66
2.965* (+/- 0.197)	-0.059* (+/- 0.008)	Acropora	194	3.046* (+/- 0.193)	-0.63* (+/- 0.007)	Acropora	54
0.330 (+/- 0.571)	0.037 (+/- 0.023)	Favia	98			Favia	2
2.873* (+/- 0.108)	-0.063* (+/- 0.004)	All Genera	835	2.099* (+/- 0.477)	-0.081 (+/- 0.018)	All Genera	122

Table A3. Comparison of our results for temperature with and without calcite restriction of ≤ 1.
*indicates significance at the 0.05 level, too few Favia samples to compute estimates.

Primary Results				No Age Restriction			
Intercept	**Temperature**	**Salinity**	**Genus**	**Intercept**	**Temperature**	**Salinity**	**Genus**
0.958* (+/- 0.305)	-0.034* (+/- 0.004)	0.031* (+/- 0.007)	Porites	0.295 (+/- 0.270)	-0.024* (+/- 0.004)	0.043* (+/- 0.006)	Porites
1.239 (+/- 1.808)	-0.057* (+/- 0.008)	0.047 (+/- 0.049)	Acropora	0.640 (+/- 0.830)	-0.044* (+/- 0.005)	0.055* (+/- 0.022)	Acropora
25.180* (+/- 8.477)	-0.059 +/- 0.040)	-0.635* (+/- 0.216)	Favia	1.754 (+/- 3.534)	-0.018 (+/- 0.021)	-0.002 (+/- 0.088)	Favia
0.659 +/- 0.389)	-0.049* (+/- 0.005)	0.053* (+/- 0.09)	All Genera	-0.315 (+/- 0.313)	-0.032* (+/- 0.004)	0.069* (+/- 0.007)	All Genera

Table A4. Comparison of our results for temperature and salinity with and without age restriction of 10,000 years.
*indicates significance at the 0.05 level

Primary Results				Calcite Restriction			
Intercept	**Temperature**	**Salinity**	**Genus**	**Intercept**	**Temperature**	**Salinity**	**Genus**
0.958* (+/- 0.305)	-0.034* (+/- 0.004)	0.031* (+/- 0.007)	Porites	2.239* (+/- 1.080)	0.013 (+/- 0.032)	-0.046 (+/- 0.048)	Porites
1.239 (+/- 1.808)	-0.057* (+/- 0.008)	0.047 (+/- 0.049)	Acropora	1.711 (+/- 1.226)	-0.061* (+/-0.007)	0.037 (+/-0.033)	Acropora
25.180* (+/- 8.477)	-0.059 (+/- 0.040)	-0.635* (+/- 0.216)	Favia				Favia
0.659 (+/- 0.389)	-0.049* (+/- 0.005)	0.053* (+/- 0.09)	All Genera	-1.956 (+/- 1.317)	-0.048* (+/- 0.018)	0.127* (+/-0.039)	All Genera

Table A5. Comparison of our results for temperature and salinity with and without calcite restriction of ≤ 1.
*indicates significance at the 0.05 level, too few Favia samples to compute estimates.

Primary Results			No Age Restriction		
Intercept	**pH**	**Genus**	**Intercept**	**pH**	**Genus**
-2.740 (+/- 1.620)	0.476* (+/- 0.200)	Porites	-4.779* (+/- 1.477)	0.730* (+/- 0.182)	Porites
74.496* (+/- 12.298)	-9.011* (+/- 1.517)	Acropora	17.196* (+/-3.296)	-1.945* (+/- 0.406)	Acropora
-1.257 (+/- 16.744)	0.310 (+/- 2.068)	Favia	13.933 (+/- 9.236)	-1.569 (+/- 1.140)	Favia
-5.113* (+/- 2.307)	0.781* (+/- 0.285)	All Genera	-3.122 (+/- 1.748)	0.539* (+/- 0.216)	All Genera

Table A6. Comparison of our results for pH with and without age restriction of 10,000 years.
*indicates significance at the 0.05 level

Primary Results			Calcite Restriction		
Intercept	**pH**	**Genus**	**Intercept**	**pH**	**Genus**
-2.740 (+/- 1.620)	0.476* (+/- 0.200)	Porites	-0.320 (+/- 5.299)	0.163 (+/- 0.658)	Porites
74.496* (+/- 12.298)	-9.011* (+/- 1.517)	Acropora	32.643* (+/- 16.425)	-3.851 (+/- 2.025)	Acropora
-1.257 (+/- 16.744)	0.310 (+/- 2.068)	Favia			Favia
-5.113* (+/- 2.307)	0.781* (+/- 0.285)	All Genera	-30.969* (+/- 5.897)	3.981* (+/- 0.730)	All Genera

Table A7. Comparison of our results for pH with and without calcite restriction of ≤ 1.
*indicates significance at the 0.05 level

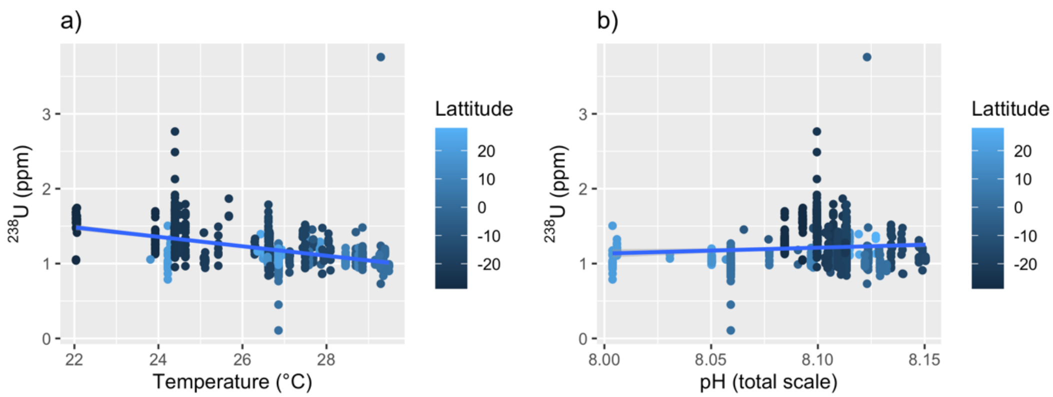

Figure A2. Relationship between ^{238}U and environmental variables of interest. As degrees in latitude increases, the color of the dots becomes lighter.

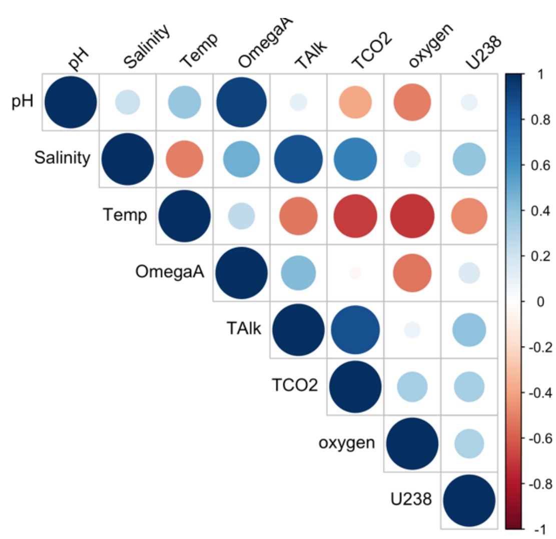

Figure A3. Correlations between all variables of interest. Larger circles indicate greater correlation. Blue circles indicate a direct relationship between the variables while red circles indicate an inverse relationship.

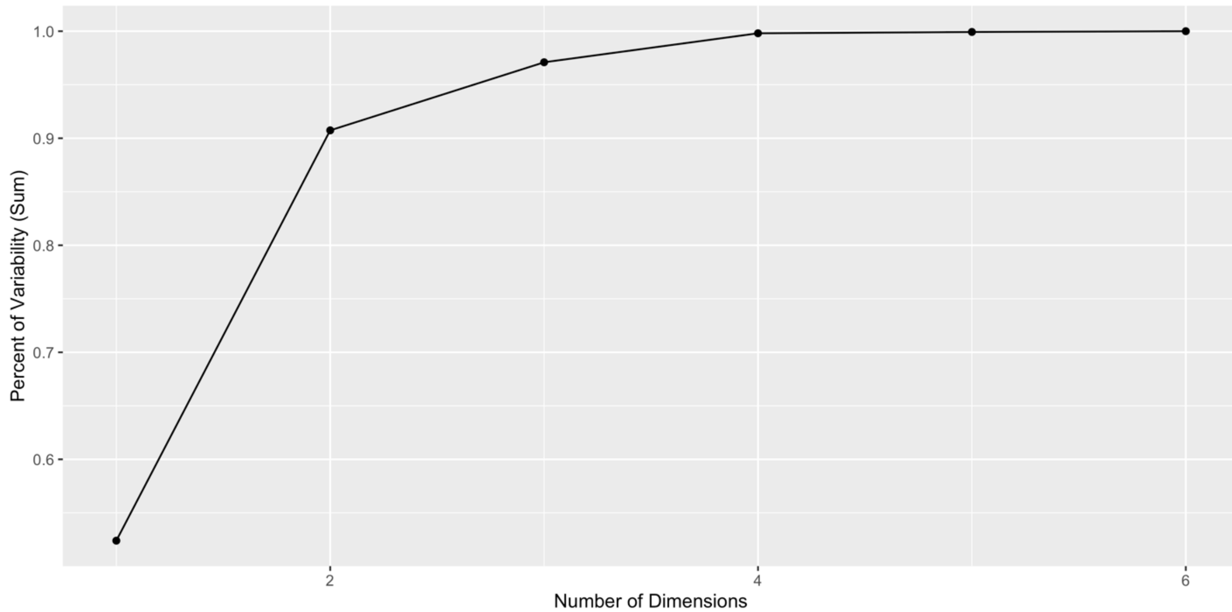

Figure A4. Sum of variability from PCA components. Here the x-axis shows the number of PCA dimensions included, and the y-axis shows the percent of variability captured by the number of PCA components. The first two components together capture 90% of the variability in the data.