1   **Supplement to:**

3   **Occupancy history influences extinction risk of fossil marine microplankton groups**

5   Isaiah E. Smith*, Ádám T. Kocsis, Wolfgang Kiessling

7   Department of Geography and Geosciences, GeoZentrum Nordbayern, Friedrich-Alexander-

8   Universität Erlangen-Nürnberg (FAU), Erlangen, Germany

10   * Corresponding author

11   Correspondence email: isaiah.em.smith@gmail.com

16

17

18

19

20

21

22

23

24

25

26

	With "gap-omission" approach (implemented in this study)	Without "gap-omission" approach
Foraminifera	3437	3039
Calcareous Nannofossils	4756	4189
Radiolarians	5522	4853
Diatoms	4767	3962

Table S1. The difference in the number of species-bin pairings with usable occupancy change and standing occupancy values for each data set binned to 1 million years, excluding occurrences occurring in bin = 1.

27

28

29

Bin size (Ma):	0.1	0.2	0.5	1.0
Foraminifera	*multiplicative*	*multiplicative*	*multiplicative*	*multiplicative*
Calcareous Nannofossils	*multiplicative*	*multiplicative*	*multiplicative*	*multiplicative*
Radiolarians	*multiplicative*	*multiplicative*	*multiplicative*	*multiplicative*
Diatoms	*multiplicative*	*multiplicative*	*multiplicative*	*multiplicative*

30 **Table S2.** The AIC-selected model formula for each group and bin size combination, on data
31 that was processed without the "gap-omission" approach. "*multiplicative*" refers to the model
32 with the formula: $ex \sim oc * ch * p$,and "*additive*" refers to the model with the formula: $ex \sim$
33 $oc + ch + p$. "*ex*" represents the binary response variable *extinction*, "*oc*" represents the
34 *occupancy* term, "*ch*" represents the *occupancy change* term, and "*p*" represents three-timer
35 sampling probability.
36

37

38

39

40

41

Simple Completeness Metric				
Bin Size (million years)	**0.1**	**0.2**	**0.5**	**1.0**
	SCM	**SMC**	**SCM**	**SCM**
Foraminifera	0.378 ± 0.007	0.470 ± 0.008	0.578 ± 0.012	0.664 ± 0.015
Calcareous Nannofossils	0.401 ± 0.006	0.493 ± 0.007	0.617 ± 0.010	0.711 ± 0.012
Radiolarians	0.372 ± 0.006	0.485 ± 0.007	0.641 ± 0.009	0.749 ± 0.011
Diatoms	0.291 ± 0.006	0.388 ± 0.007	0.530 ± 0.010	0.640 ± 0.012

Table S3. Simple completeness metric (SCM, Benton 1987) calculations for each data set at each of the four examined bin sizes. Shown with standard error.

Bin size (Ma):	**0.1**	**0.2**	**0.5**	**1.0**
Foraminifera	*multiplicative*	*multiplicative*	*multiplicative*	*multiplicative*
Calcareous Nannofossils	*multiplicative*	*multiplicative*	*multiplicative*	*multiplicative*
Radiolarians	*additive*	*additive*	*multiplicative*	*multiplicative*
Diatoms	*multiplicative*	*multiplicative*	*multiplicative*	*additive*

Table S4. The AIC-selected model formula for each group and bin size combination, on extinct-only data. "Multiplicative" refers to the three-term, multiplicative model. "Additive" refers to the three-term additive model.

Pearson correlation between total number of occupied Longhurst provinces and total number of occupied cells (corrected for autocorrelation)				
Bin Size (million years)	0.1	0.2	0.5	1.0
Foraminifera	0.839	0.841	0.875	0.898
Calcareous Nannofossils	0.878	0.857	0.892	0.915
Radiolarians	0.713	0.736	0.792	0.822
Diatoms	0.810	0.809	0.826	0.860

59 **Table S5.** Pearson correlation values between the total number of occupied Longhurst
60 provinces and the total number of occupied cells for each taxonomic group at each bin size.
61 All correlation coefficients were highly significant ($p < 10^{-5}$).
62

63

64

65
66
67

68

69

70

71

72

73

74

75

76

77

78

79

80

Foraminifera

81

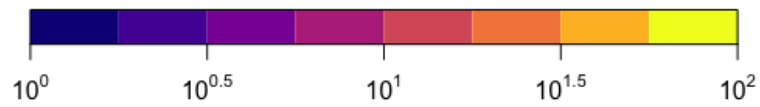
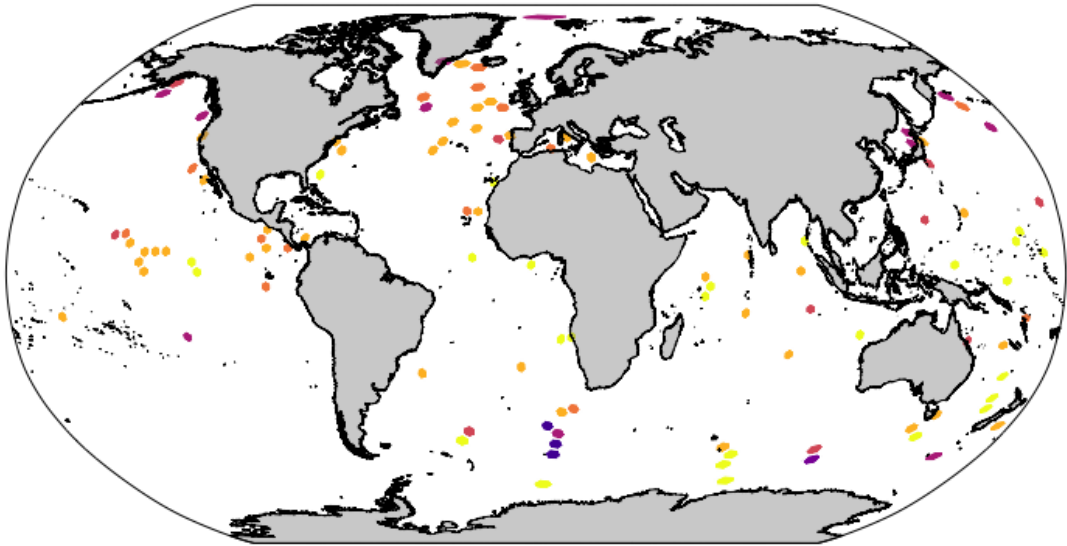
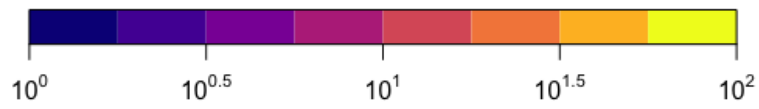
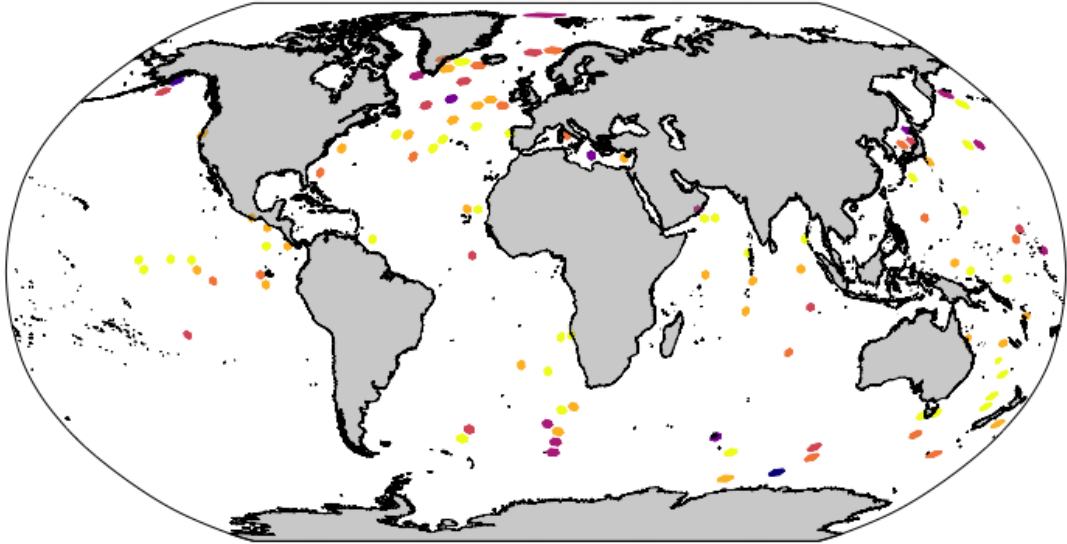



Figure S1A. Number of foraminifera occurrences in each geographic cell. Cells assigned according to each occurrence's modern geographic coordinates. White areas correspond to cells without records.

85

Calcareous Nannofossils

86

87   **Figure S1B.** Number of calcareous nannofossil occurrences in each geographic cell. Cells
88   assigned according to each occurrence's modern geographic coordinates. White areas
89   correspond to cells without records.

Radiolarians

10^0 $10^{0.5}$ 10^1 $10^{1.5}$ 10^2

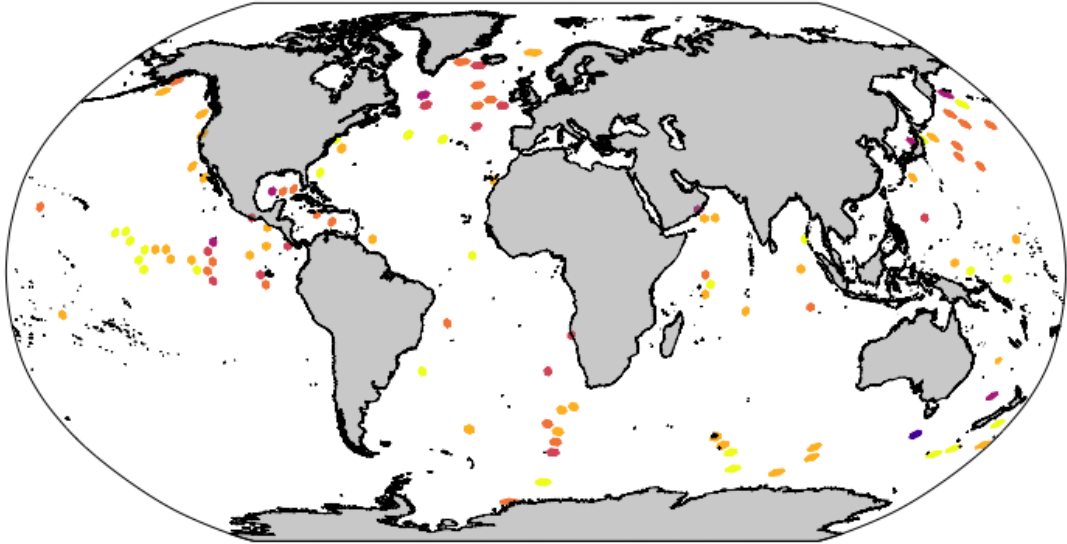
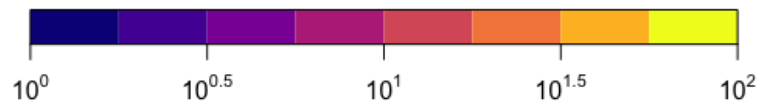



Figure S1C. Number of radiolarian occurrences in each geographic cell. Cells assigned according to each occurrence's modern geographic coordinates. White areas correspond to cells without records.

Diatoms

95

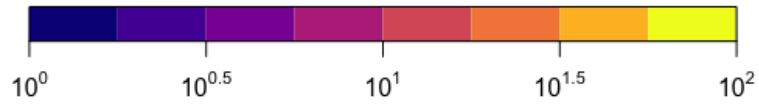
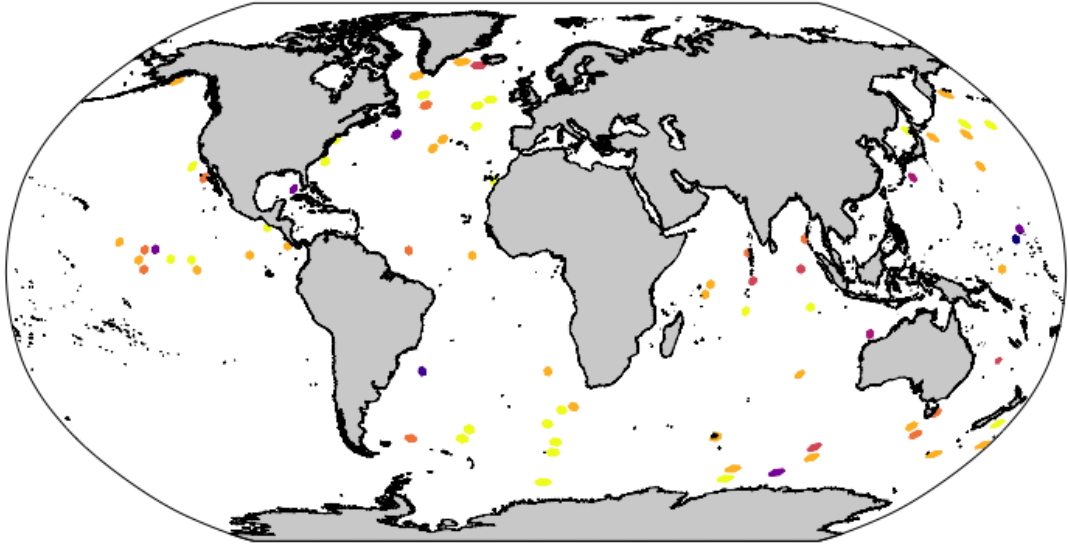
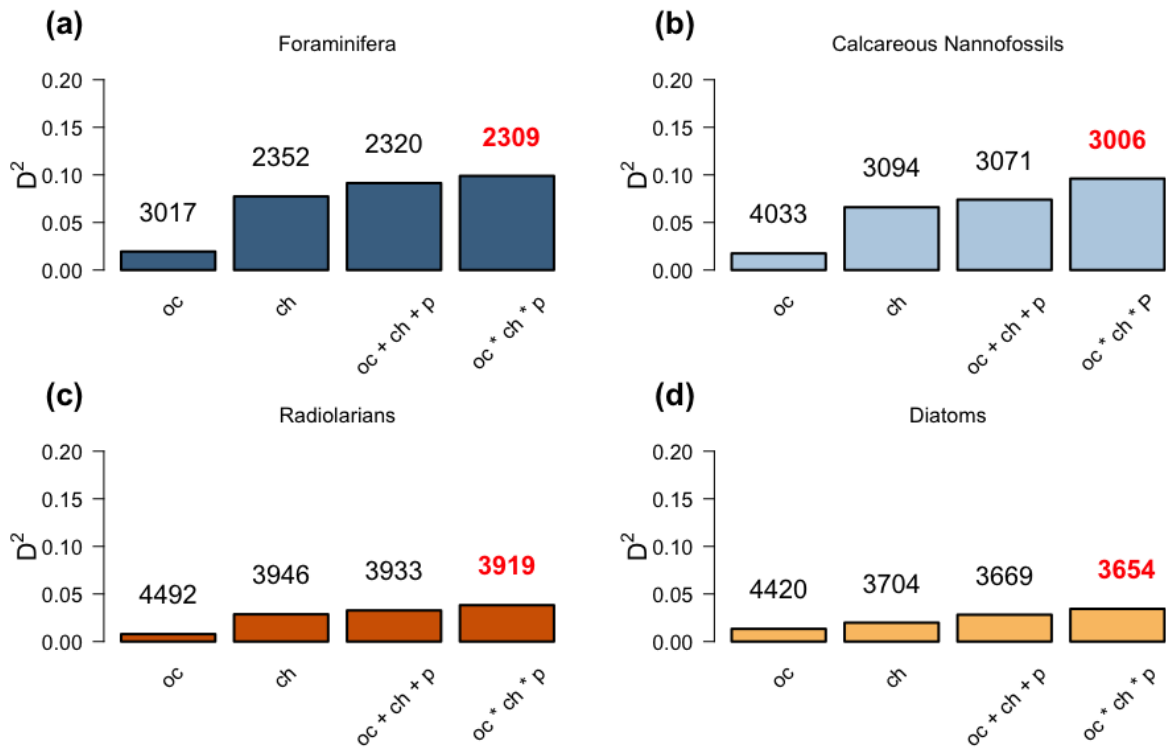
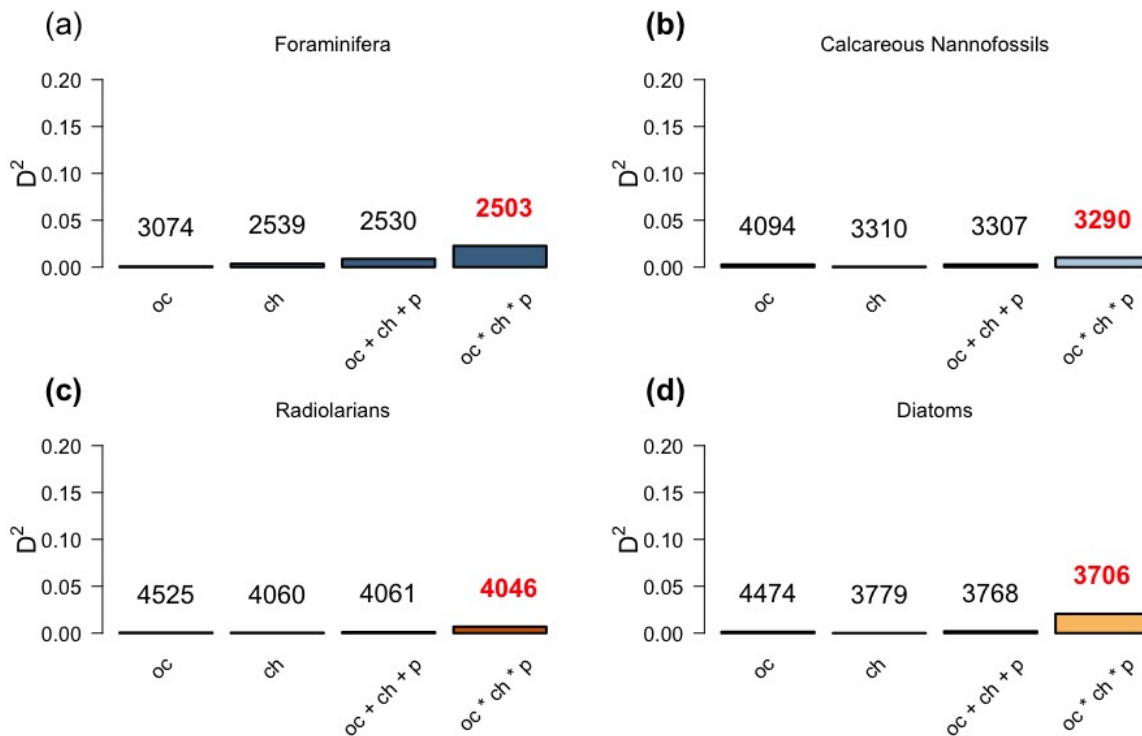
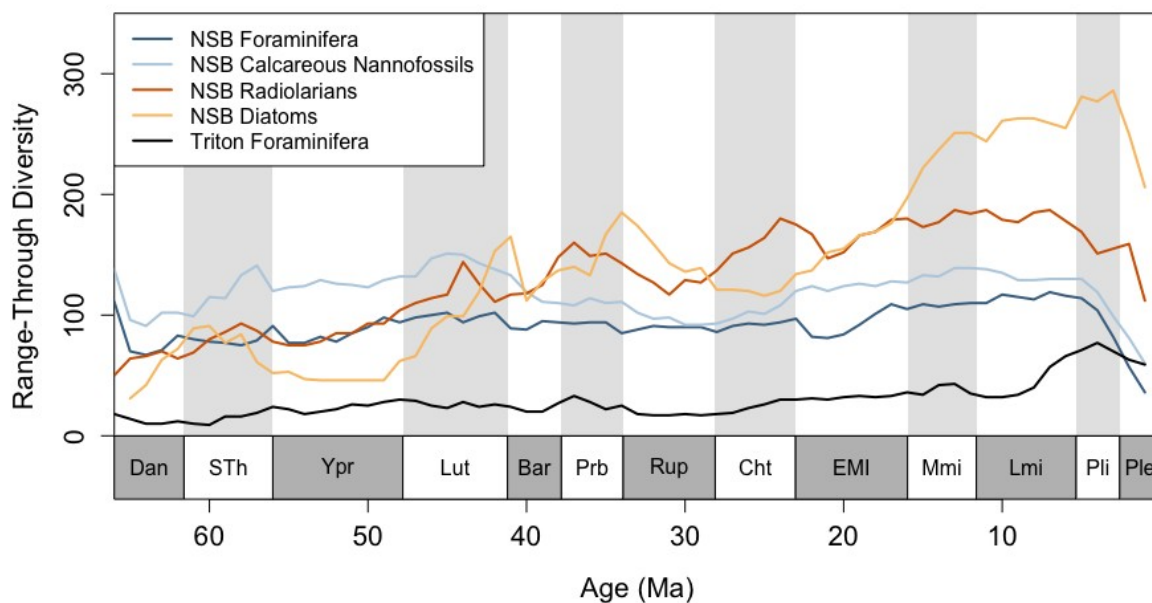
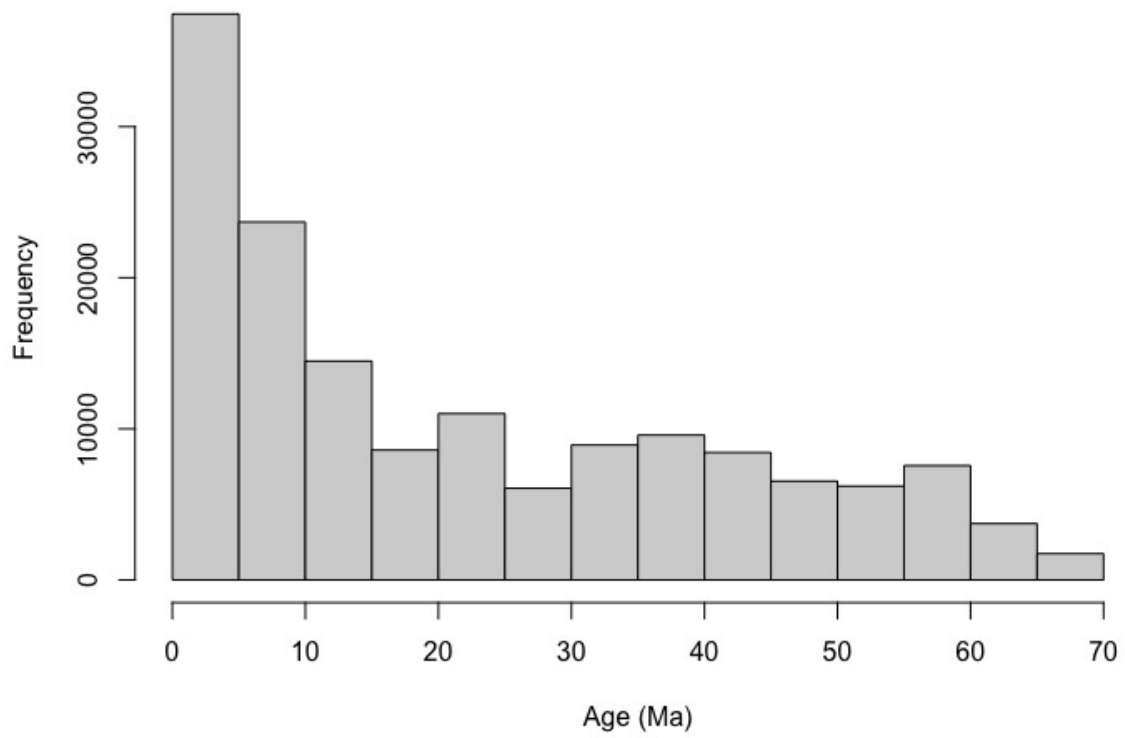
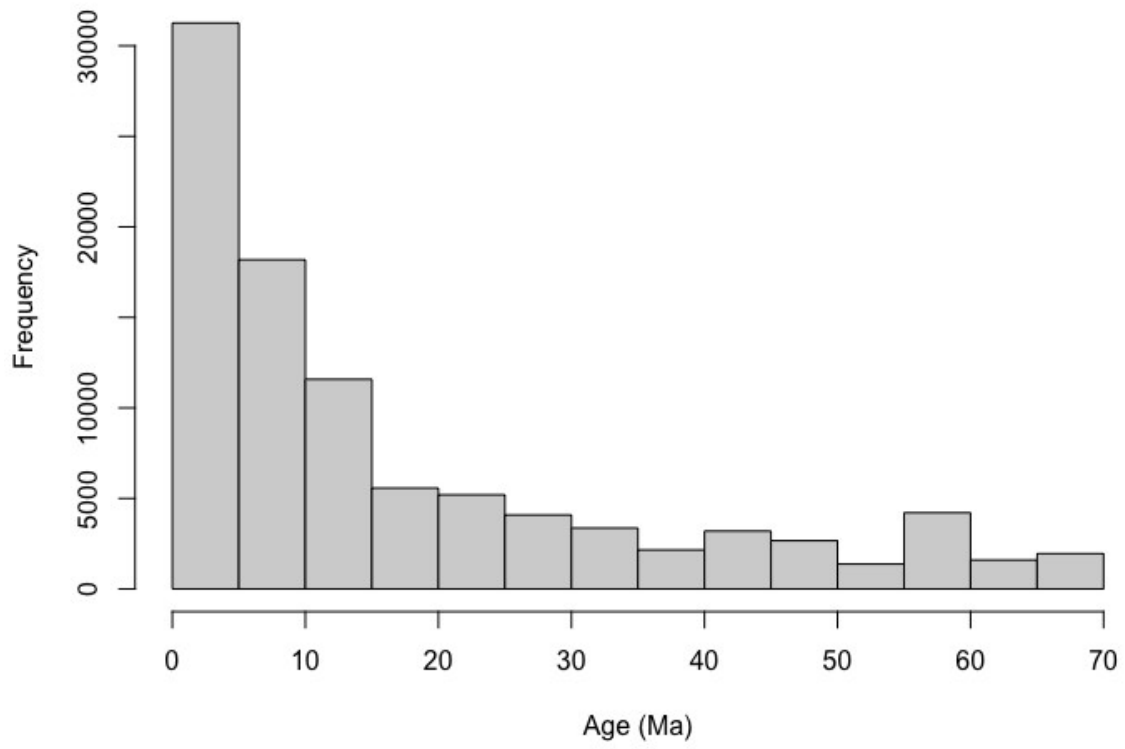
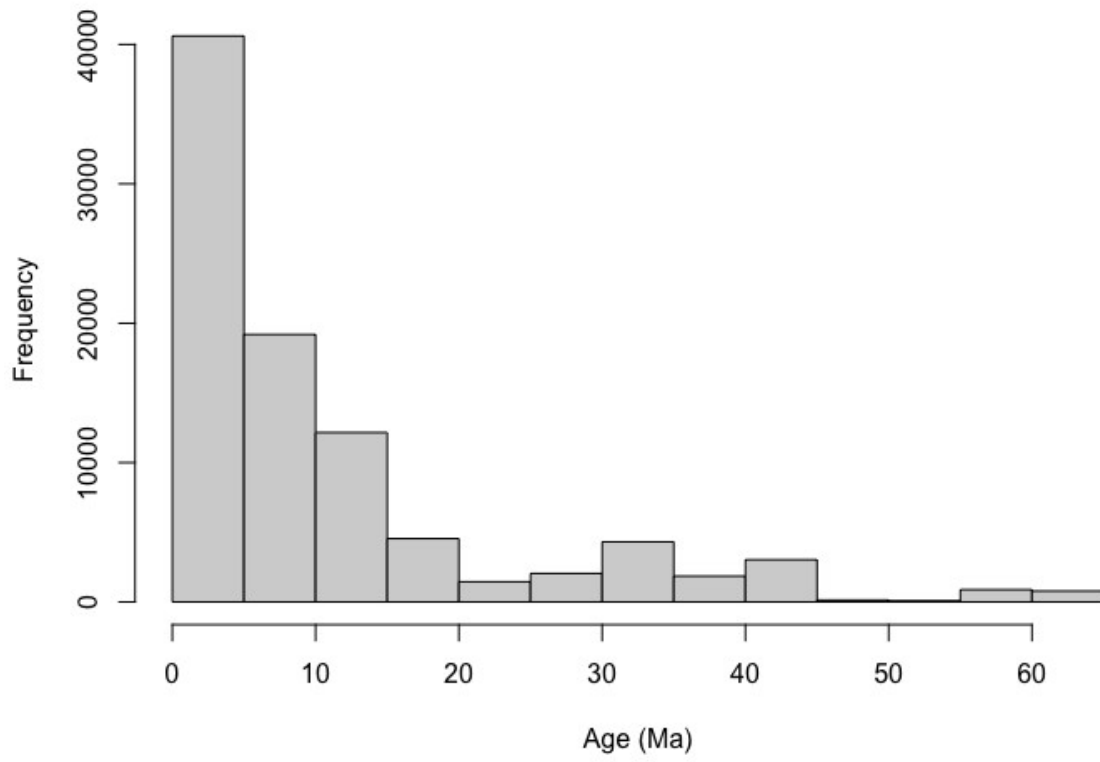
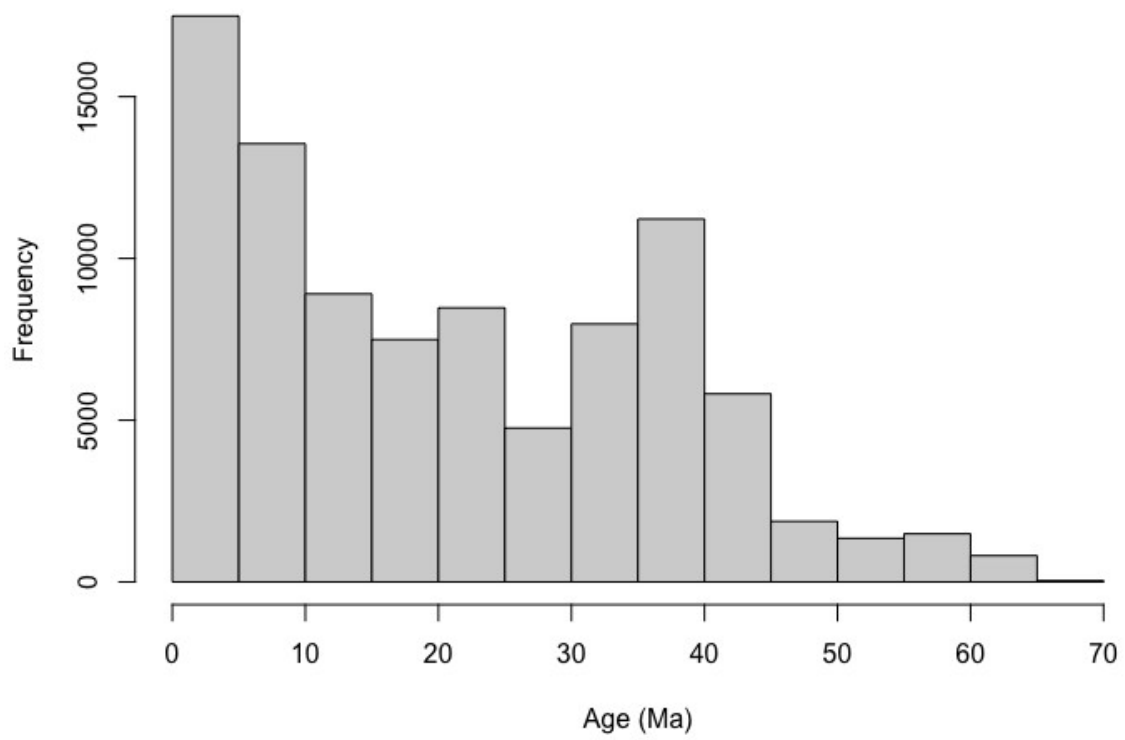
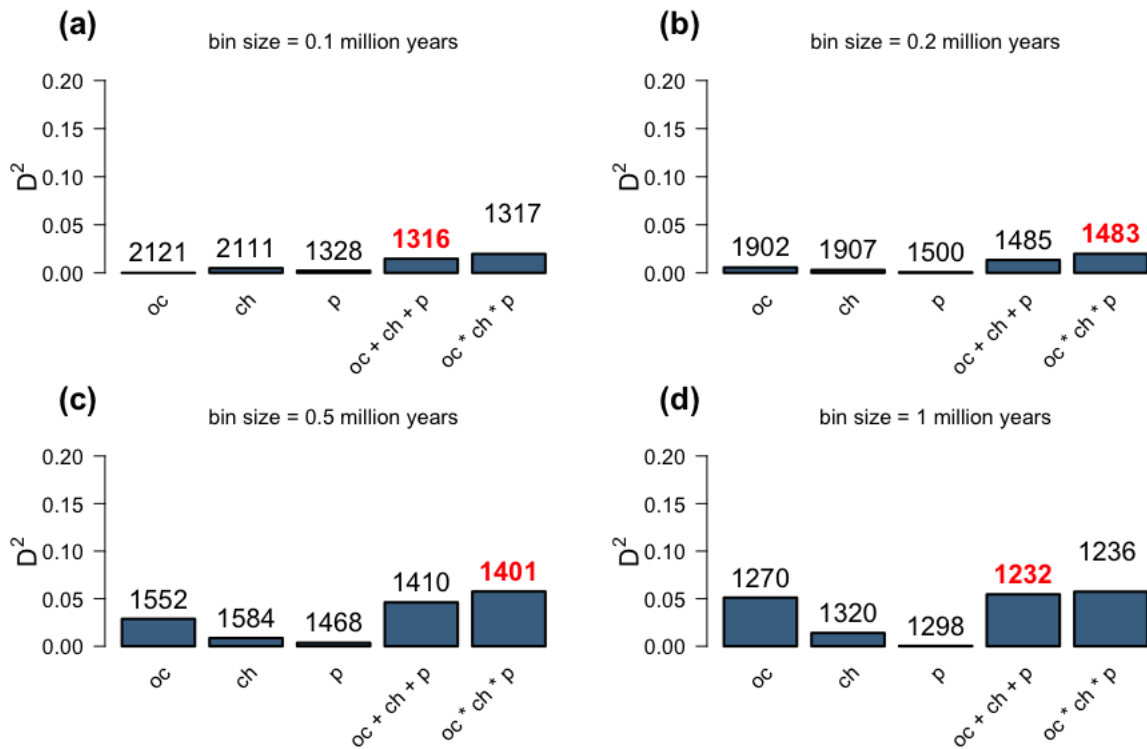



Figure S1D. Number of diatom occurrences in each geographic cell. Cells assigned according to each occurrence's modern geographic coordinates. White areas correspond to cells without records.

99 **Figure S2.** D^2 values (bars) and AIC values (numbers displayed above each bar) for each
100 examined model formula using proportional occupancy of Longhurst provinces and change
101 in proportional occupancy of Longhurst provinces. Shown for each of the NSB data sets
102 with bin size = 1 million years. The lowest AIC value for each data set is shown in red.
103
104
105

106
107 **Figure S3.** D^2 values (bars) and AIC values (numbers displayed above each bar) for each
108 examined model formula using latitudinal range and change in latitudinal range. Shown for
109 each of the NSB data sets with bin size = 1 million years. The lowest AIC value for each
110 data set is shown in red.
111

112

113
114 **Figure S4.** Range through diversity for each of the four NSB datasets and the Triton dataset,
115 using a bin size of 1.0 million years. Notice how the Triton dataset has consistently lower
116 diversity than the other four datasets. Also note the decrease in diversity approaching the
117 present, suggesting sampling issues at the tops of drill cores.
118
119
120
121
122

123
124 **Figure S5A.** Histogram of occurrence frequency of NSB calcareous nannofossils.
125
126
127
128

129
130   **Figure S5B.** Histogram of occurrence frequency of NSB foraminifera.
131
132
133
134

135
136 **Figure S5C.** Histogram of occurrence frequency of NSB diatoms.
137
138
139
140

141
142 **Figure S5D.** Histogram of occurrence frequency of NSB radiolarians.
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159

160
161 **Figure S6.** D^2 values (bars) and AIC values (numbers displayed above each bar) for each
162 examined model formula using the Triton dataset at each of the four analyzed bin sizes. The
163 lowest AIC value for each data set is shown in red.
164
165
166
167
168
169
170
171
172