bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Community structure and sample size affect estimates of predation intensity and prey selection: A model-based validation

Madhura Bhattacherjee, Devapriya Chattopadhyay*

RRH: COMMUNITY STRUCTURE AFFECTS INFERRED PREDATION

LRH: BHATTACHERJEE AND CHATTOPADHYAY

Keyword: Drilling predation, durophagous predation, evenness, sampling, predation intensity

* Corresponding author

Email: devapriya@iiserpune.ac.in

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract. –

Predation estimates inferred from the preserved records of predation traces are essential in evaluating the evolutionary effect of ecological interactions. It is, however, rarely investigated how sampling intensity and community composition of an assemblage influence the reliability of these measures.

Using a resampling technique, we evaluated the effect of evenness and sampling intensity of a community on the inferred predation estimates. We theoretically simulated model communities representing different levels of evenness (E_T), predation intensity (PI_T), and predatory behavior (selective, non-selective). These communities were resampled without replacement; we noted the variation in the inferred predation intensity ($PI_{T.inf}$) and the number of prey species ($S_{prey.inf}$) during resampling. Our results demonstrate that E_T does not influence $PI_{T.inf}$ for non-selective predation. However, communities with highly selective predation are sensitive to evenness and sampling intensity; $PI_{T.inf}$ of these assemblages can substantially deviate from the actual value. $S_{prey.inf}$ is also influenced by E_T, predation selectivity, and PI_T. For non-selective predation, sampling intensity heavily influences communities with low E_T and low PI_T; $S_{prey.inf}$ is underrepresented at smaller sample size. For prey-abundance guides selective predation, the sensitivity depends on the nature of selection; for communities of low E_T and PI_T where rare species are attacked preferentially, the $S_{prey.inf}$ deviates significantly at small sample size. We proposed a post-facto standardization method for comparing predation estimates of discrete communities. We validated its utility using the published predation data of the Plio-Pleistocene molluscan fossil assemblage. This approach provides critical insight into the biological reliability of predator estimates across time and space.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Madhura Bhattacherjee, Department of Earth and Climate Science, IISER Pune, Dr. Homi Bhaba Road, Pashan, Pune-411008, INDIA

Devapriya Chattopadhyay, Department of Earth and Climate Science, IISER Pune, Dr. Homi Bhaba Road, Pashan, Pune-411008, INDIA, Email: devapriya@iiserpune.ac.in

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

The role of predation as a major driving force of evolution in marine ecosystems has been a common theme of study (Paine, 1966; Vermeij, 1977; Signor and Brett, 1984; Langerhans, 2007; Stanley, 2008). The relationship between the prey and predator is complex in theoretical and inferential terms (Abrams, 1989; Leighton, 2002). It is important to note that efforts to evaluate the evolutionary effects of predation relies on the correct measurement of aspects of predation such as intensity, selectivity of prey. The accurate estimation of predation measures is, therefore, of primary importance.

For establishing predation events and inferring predation intensities, ecological studies use direct observations or indirect measures such as compositional characterization of digested food and fecal matter (Nilsen et al., 2012; Pringle et al., 2019). Although it is possible to recover direct observational evidence of predation events in past ecosystems by studying “caught-in-the-act” occurrences, paleoecological studies primarily rely on preserved predation traces (Klompmaker et al., 2019). Traces, such as drill holes and repair scars recording the predation attempts on the prey’s hard shells, are some of the best quantifiable proxies for inferring predation intensity and prey selection (Kowalewski, 2002). The frequency of these traces (repair scar frequency (RF) and drilling frequency (DF)) are often used to evaluate the evolutionary impact of predation in deep time (Vermeij, 1987; Kelley and Hansen, 1996).

Inferences about interactions from predation traces have their limitations. The implicit assumption for such interpretation is that other processes do not alter the quantitative data provided by predation traces (e.g., the overall frequency of traces, distribution of traces across prey species, and prey size classes). It is recognized that biases

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

introduced through taphonomy may influence the biological reliability of these measures, including frequency (Roy et al. 1994; Zuschin and Stanton, 2001, Chattopadhyay et al., 2013a, b, Sarkar et al., 2021), prey selection (Sime and Kelley 2016; Forcino et al. 2017), and size selection (Kowalewski, 2002). However, the effect of sampling on inferred predation measures remains largely unexplored.

A few studies have examined the impact of size-specific sampling on inferred predation intensities (Chattopadhyay et al., 2016; Hattori et al., 2014; Hausmann et al., 2018; Kowalewski and Hoffmeister, 2003; Ottens et al., 2012; Visaggi and Kelley, 2015). Smith et al. (2021) demonstrated the effects of overdispersion and zero inflation (e.g., the poor model fits) using count data of predation traces. Aspects of a specific community such as evenness, selectivity of predation, and sampling intensity may influence the estimates of predation intensity. This becomes crucial for studies that combine predation data from discrete samples and reconstruct temporal/spatial changes in predation patterns.

Using theoretical simulation based on a resampling technique, we developed a methodological framework to understand the effect of community evenness, sampling intensity, and nature of predation selectivity on the inferred predation estimates. We attempted to estimate these effects on the inferred predation intensity and the number of prey species. We also proposed a method of post-facto standardization and validated our approach using predation data from four Plio-Pleistocene fossil assemblages of Florida.

Materials and method

Indices for predation intensity and community structure:

Predation intensity of total assemblage (PI_T) is calculated as

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$PI_T = N_P/N$

Where,

N_P = Number of individuals with predation mark

N = Total number of individuals in the assemblage

The total number of prey species is S_{prey}. PI_{prey} denotes predation intensity in the species that have been attacked.

The evenness of an assemblage (E_T) is calculated as

$E_T = H/ \ln(S_T)$

Where,

H = Shannon's diversity index

S_T = Total number of species in the assemblage

Model assemblages:

We created 30 hypothetical model assemblages, each with 30 species and 3000 individuals with varying evenness, predation intensity, and prey preferences (Table 1). Each model assemblage had a unique combination of evenness, predation intensity, and prey preference. The evenness in these models ranged from a theoretical minimum of 0.1 to a theoretical maximum of 1. Model assemblages with maximum evenness of one had 100 individuals for 30 species. Assemblages with intermediate evenness of 0.7 had five common species with 500 individuals each and 25 rare species with 20 individuals each (Table 1). Assemblages with low evenness of 0.5 had 910 individuals in each of the three common species and ten individuals in each of the 27 rare species. Assemblages with a very low

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

evenness of 0.2, there is only one common species with 2710 individuals, and the remaining 29 rare species consists of 10 individuals each.

The predation intensity of the total assemblage (PI_T) was categorized into three levels: low (0.2), medium (0.5), and high (0.8) (Table 1). A certain number of individuals from specific species would be considered prey with predation marks as dictated by the (PI_T). The prey-preference of the predator can either be non-selective or selective. In the case of non-selective predation (Case 1), all species have an equal probability of being attacked irrespective of their abundance (Fig 1). Selective predation represents assemblages where prey species have an unequal chance of being attacked depending on their abundance. In model assemblages with selective predation, we constructed two cases; the predator can either attack the common species (Case 2) (Fig 2) or the rare species (Case 3) (Fig 3). Selective predation has not been considered for assemblages with maximum evenness because all the species are equally abundant and share an equal probability of attack.

Simulation design

We performed a simulation to evaluate the effect of sample size on inferred predation intensity ($PI_{T.inf}$) and the number of prey species ($S_{prey.inf}$) for all the model assemblages. In the simulation, 100 individuals were drawn randomly from a model assemblage. The number of attacked individuals (N_P) and the number of prey species (S_{prey}) represented by the attacked individuals were counted in those 100 individuals. Inferred predation intensity ($PI_{T.inf}$) for the drawn sample is calculated as a ratio of the number of attacked individuals and the total number of individuals (i.e., 100 in the first draw). The exact process is repeated 30 times without replacement until all the individuals from the

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

assemblage are sampled. This entire process was iterated 1000 times. The mean and standard deviation are calculated for inferred predation intensity ($PI_{T.inf}$) and prey species richness ($S_{prey.inf}$) for these 1000 iterations for a model assemblage. Deviation of predation intensity (Dev_{PI}) is calculated as the difference between PI_T and $PI_{T.inf}$ for an assemblage. Similarly, the difference between S_{prey} and $S_{prey.inf}$ is taken as the deviation of prey species richness (Dev_S). The same technique is applied for all the model assemblages.

Predation dataset:

We used published data on predation records of molluscs from four Pleistocene localities in Florida (Chattopadhyay and Baumiller, 2010) for validating the proposed technique. The dataset consists of abundance, drilling frequency, and repair scar frequency of 14 molluscan species. We drew samples without replacement from each locality with increasing sample size. The sample size for each draw was a hundred until the last draw; in the last draw, where the remaining sample size is less than 200, all are drawn. For Punta Gorda (total=2418 individuals), 100 individuals were drawn 23 times, and 118 individuals were drawn for the last (24th) draw. A similar procedure is followed for Miami Canal (total = 4794 individuals), Mc Queens pit (total=659 individuals), and Chiquita (total=894 individuals).

All simulations and statistical analyses were performed in R (R Core Development Team, 2012).

Results

Inferred predation intensity

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The inferred predation intensity ($PI_{T.inf}$) may vary substantially from the actual value of overall predation intensity (PI_T) and predation intensity of prey groups (PI_{prey}), especially at smaller sample sizes (Fig 4). For non-selective predation (Case 1), Dev_{PI} is not affected by the evenness; however, the deviation is high at a smaller sample size and converges to the actual value with increasing sample size (Fig 5). Evenness influences inferred predation intensity ($PI_{T.inf}$) when the predation is non-selective (Case 2 and 3) (Fig 4). When the common species are preferentially attacked (Case 2), Dev_{PI} is low for communities with lower evenness and low original predation intensity (PI_T) (Fig 6). Communities with higher evenness showed high Dev_{PI}. This deviation does not change even at a higher sample size except for the community with the lowest evenness (E_T=0.2) and predation intensity (PI_T=0.2) (Fig 6). When rare species are attacked (Case 3), Dev_{PI} is high for communities with higher evenness and high PI_T (Fig 7). Communities with higher evenness showed lower $PI_{T.inf}$ values compared to PI_T. The deviation, however, is lower in comparison to comparable communities in Case 2.

Inferred number of prey species

The inferred number of prey species ($S_{prey.inf}$) follows a rarefaction curve where $S_{prey.inf}$ increases with increasing sample size before plateauing and converging to the actual value of S_{prey} (Fig 8). In the case of non-selective predation (Case 1), the Dev_S decreases with increasing sample size. Dev_S decreases with increasing evenness at a given sample size and has lesser sensitivity to sample size when evenness is high. Dev_S approaches zero at a smaller sample size in communities with higher evenness (Fig 9). When common species are preyed upon (Case 2), Dev_S does not reflect any sensitivity to the sample size (Fig 10).

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

This is due to the low value of S_{prey} that converges to its actual value within the first few draws.

However, when the rare species are attacked (Case 3), $S_{prey.inf}$ is highly sensitive to the sample size. Communities with higher evenness tend to show lesser variation in Dev_S with sample size than others (Fig 11-12). The overall Dev_S in all the communities is more in lower PI_T.

Inferred predation estimates from Florida

The assemblages from the four localities of Florida are different in terms of their evenness and sample size (Table 2). Except for Miami canal, all the localities show a case of non-selective predation represented by the lack of correlation between the relative abundance of the prey and prey-specific predation intensity (PI_{prey}) for drilling and durophagous predation (Table 3). In Miami canal, the significant positive correlation implies that the abundant prey species are preferentially attacked – a scenario described as Case 2 in our model.

For inferred predation intensity ($PI_{T.inf}$), there is substantial overlap between three localities (Punta Gorda, Miami Canal, and Mc Queens pit) for both drilling and durophagy (Fig 13).

For inferred prey species richness ($S_{prey.inf}$), the assemblages show slightly different patterns between drilling and durophagous predation. For drilling predation, all the assemblages show a substantial overlap (Fig 13A, C, E). The durophagous predation record, however, shows a separation between communities with low evenness (Punta Gorda) and high-evenness (Mc Queens pit, Chiquita) (Fig 13 B, D, F).

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

Paleontological research on predation has expanded rapidly in scope, methods, and goals over the years. Various studies have focused primarily on identifying the evidence of fossil records of predation and their interpretation; such records have also been used for testing ecological and evolutionary hypotheses. In contrast, studies exploring the reliability of the methods to quantify predation intensity have been limited. While evaluating predation intensity, the intensity matrices depend more on the number of individuals of a community and less on the community structure. Our model provides a theoretical framework to estimate inferred predation intensities for various community structures and sampling intensities. It highlighted the importance of these factors in influencing the predation estimates of fossil and recent assemblages; it also demonstrates why it is necessary to develop a methodological framework of sample standardization before comparing predation estimates of assemblages separated by time and space.

Effect on the inferred intensity

Our simulation results show that communities' evenness does not significantly change the inferred predation intensity when random encounters between predator and prey guide predation. It is, however, uncommon to find predation events to be completely random in the natural world. The inferred predation intensity may deviate significantly from the original predation intensity if the predation is selective.

Following the considerations of optimal foraging theory, two aspects make the predation selective. The first is the relative ease with which a predator encounters a prey – a factor primarily guided by the relative abundance of the prey. With the increasing dominance of a

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

prey species, the probability of encounter increases, and foraging time decreases; this increases the net energy gain of the predator. The second aspect is the traits (morphological, ecological, behavioral) of the prey that dictates the net energy gain of the predator. The final selection by the predator is often a combination of these factors. A higher attack rate may be found in an abundant prey species due to its higher encounter rate despite its high anti-predatory defense than a rarer species (Johnson, 1980). This would lead to scenarios similar to Case 2, where the inferred predation intensity of low-evenness communities would be higher than the actual predation intensity. This inflation is because of the over-representation of common species in smaller samples that are primarily attacked.

Most often than not, the encounter frequency does not finally dictate the attack frequency, and the selection of prey is guided by the prey traits such as size (Chattopadhyay et al., 2014, 2015, 2020). These would be similar to Case 3, where the most dominant groups are not preyed upon. The inferred predation intensity of low-evenness communities would be lower than the actual predation intensity. This apparent drop in predation intensity is because of the lack of representation of rare species in smaller samples that are never attacked. It is especially problematic because the deviation between PI_T and $PI_{T.inf}$ is substantial for all evenness.

Effect on inferred selectivity

The structure of the community heavily controls the inferred number of prey species. Communities with low evenness deviate significantly from the original prey species and yield fewer inferred prey species even when the predation is non-selective (Case 1). This may lead to the development of an artificial selectivity primarily driven by the preferential

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

counting of the dominant species and not by the biological preference demonstrated by the predators. Therefore, any community with low evenness suffers from the high probability of underrepresenting the number of prey species. The deviation from the true prey-species richness is higher for smaller sample size and lower intensity of predation. Communities with higher predation intensity will provide the true prey-species richness at a smaller sample size than communities with lower predation intensity. Selective predation (as indicated by Case 2 and 3) also creates similar deviations.

The sensitivity of inferred prey species richness on sample size, evenness, and original predation selectivity makes the comparison of prey species richness in spatially or temporally distinct assemblages somewhat unreliable unless they are normalized. This is especially important when comparing predation estimates from assemblages representing different time-bins or environments likely to show varying diversity/evenness.

Proposed protocol of post-facto standardization of predation data

The following protocol may be followed to compare predation intensity and selectivity of spatially/temporally distinct assemblages to avoid misinterpretation. Using the protocol described for generating Figure 4, all the assemblages need to be plotted and compared to understand the sensitivity of the inferred predation intensity ($PI_{T.inf}$) on sample size. If this plot shows a significant overlap between different assemblages, they should be considered to represent the same predation intensity (PI_T), and the difference in sampling size explains the difference in the absolute value. To understand the sensitivity of the inferred prey-species richness ($S_{prey.inf}$) on sample size, the protocol described for generating Figure 5 needs to be followed for all the assemblages and compared. If this plot shows a significant overlap

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

between different assemblages, they should be considered to represent the same prey-species richness (S_{prey}), and the difference in sampling size explains the difference in the absolute value. The assemblage with the smallest sample size can be considered as a reference. The distribution of inferred predation intensity ($PI_{T.inf}$) and prey-species richness ($S_{prey.inf}$) at that sample size should be compared for all the assemblages using Kolmogorov-Smirnov (K-S) test. Unless the difference is significant, we cannot reject the possibility of inadequate sampling generating the difference. We should select the assemblages from a pairwise comparison of all the assemblages that yield statistical significance. We should choose assemblages with a larger sample size among the pairs producing without statistical significance. Using those selected assemblages, we should repeat the entire process at the smallest sample size of the new subset of the original assemblage. We should continue this iteration till all the pairs show significant differences. This difference cannot be explained by sampling and reflects the biological difference. Estimating PI_{prey} is difficult, especially for cases where rare species are attacked; excluding species without any predation trace while calculating PI_T may give us some insight.

Paleontological case study

The assemblages from the four localities of Florida have been used for interpreting the relationship between durophagy and drilling predation (Chattopadhyay and Baumiller, 2010). However, the study's conclusions did not consider sample size or community structure. The assemblages are of these localities are pretty different in terms of their evenness and sample size (Table 2). A selective nature of predation where common prey species are preferentially attacked is observed for both predations (durophagy and drilling)

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

only in Miami Canal. For inferred predation intensity ($PI_{T.inf}$), the sample size-standardized resampling protocol (described before) shows a significant difference (Table 4). This implies that the difference across assemblages cannot be explained by the sampling intensity or the evenness of the assemblage. Similarly, the sample size-standardized resampling protocol (described before) yields a significant difference for all the assemblages for inferred prey species richness ($S_{prey.inf}$) (Table 4). This implies that the difference in the observed prey-species richness across assemblages cannot be explained by sample intensity/evenness.

It is important to recognize that a number of factors played a role in this particular case that made these assemblages less susceptible to community evenness and sampling intensity. Because three localities (Punta Gorda, Mc Queens pit, Chiquita) are showing non-selective predation, they are less likely to be affected by sample size. Moreover, they have medium to high evenness that makes them less sensitive to sample size. Miami canal shows evidence of selective predation where abundant species are preferentially attacked. The expected deviation in predation intensity of the assemblage of Miami canal, characterized by low evenness (0.31), low predation intensity ($PI_T<0.2$) should follow Fig 6A that shows a low value of Dev_{PI} at large sample size. Because Miami canal has the largest sample size among the localities, makes it less likely to be affected by these factors. The expected deviation in inferred prey species richness (Dev_S) of the assemblage of Miami canal should follow Fig 10A that shows least sensitivity to sample size. Hence, the observed $S_{prey.drill}$ and $S_{prey.repair}$ are least likely to be affected by these factors.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Implications

The fossil record of predation has shaped our understanding of how the nature of biotic interaction changed over time and its role as an evolutionary mechanism. Preserved traces, such as drill holes and repair scars, are some of the best quantifiable proxies of predation and they are often used to assess the evolutionary impact of predation in deep time (Vermeij, 1987; Kelley and Hansen, 1996, Huntley and Kowalewski, 2007). Studies aiming to evaluate the predation trend through time, however, are often forced to use predation data from discrete assemblages that differ in sample size, inherent community evenness, and the type of predation selectivity. Our study demonstrates the effect of such factors on the inferred predation intensity and the recognized prey richness. Comparison between temporally separated collections, such as Paleozoic and Cenozoic predation records that are known to be different in the sample size (and probably predatory behavior), are susceptible to such factors. Our proposed method of post-facto standardization will be essential for such comparisons and to establish the true nature of biotic interaction through time.

Conclusions

The effect of community structure and sampling intensity on the inferred predation estimates is rarely explored. Using a resampling technique, our study demonstrates the impact of these aspects on the estimates of predation intensity and the number of prey species. Our results show that the communities with highly selective predation are the most sensitive to sampling intensity, and the inferred predation intensity of these assemblages can substantially deviate from the actual value. In contrast, predation intensity for non-selective predation tends to be unaffected by sampling intensity. Inferred prey-species richness is also

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

influenced by the nature of community evenness, predation selectivity, and actual predation intensity. For non-selective predation, communities with low evenness and low predation intensity are highly sensitive to sample size. The inferred prey-species richness can be underrepresented significantly at smaller sample size. For selective predation, the sensitivity depends on the nature of selection. The inferred prey-species richness deviates significantly when rare species are attacked preferentially. Our study also provides a framework of post-facto standardization of the predation data to remove the effect of sample size/evenness during comparison. The proposed method will be essential for temporal comparison of pre- and post-Cretaceous assemblages as they are often characterized by a difference in sample size/selectivity.

Acknowledgments

This project was supported by the SERB core research grant (CRG/2018/002604). The IISER Pune Doctoral fellowship supported MB.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Abrams, P. A. 1989. The evolution of rates of successful and unsuccessful predation. Evolutionary Ecology 3:157-171.

Chattopadhyay, D., A. Rathie, and A. Das. 2013a. The effect of morphology on postmortem transportation of bivalves and its taphonomic implications. Palaios 28:203-209.

Chattopadhyay, D., A. Rathie, D. J. Miller, and T. K. Baumiller. 2013b. Hydrodynamic effects of drill holes on postmortem transportation of bivalve shells and its taphonomic implications. Palaios 28:875-884.

Chattopadhyay, D., and T. K. Baumiller. 2010. Effect of durophagy on drilling predation: a case study of Cenozoic molluscs from North America. Historical Biology 22:367-379.

Chattopadhyay, D., D. Sarkar, S. Dutta, and S. R. Prasanjit. 2014. What controls cannibalism in drilling gastropods? A case study on Natica tigrina. Palaeogeography, Palaeoclimatology, Palaeoecology 410:126-133.

Chattopadhyay, D., M. Zuschin, and A. Tomašových. 2015. How effective are ecological traits against drilling predation? Insights from recent bivalve assemblages of the northern Red Sea. Palaeogeography, Palaeoclimatology, Palaeoecology 440:659-670.

Chattopadhyay, D., M. Zuschin, S. Dominici, and J. A. Sawyer. 2016. Patterns of drilling predation in relation to stratigraphy, locality and sieve size: Insights from the Eocene molluscan fauna of the Paris Basin. Palaeogeography, Palaeoclimatology, Palaeoecology 459:86-98.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Chattopadhyay, D., V. G. S. Kella, and D. Chattopadhyay. 2020. Effectiveness of small size against drilling predation: Insights from lower Miocene faunal assemblage of Quilon Limestone, India. Palaeogeography, Palaeoclimatology, Palaeoecology 551:109742.

Hattori, K. E., P. H. Kelley, G. P. Dietl, N. O. Moore, S. L. Simpson, A. M. Zappulla, K. J. Ottens, and C. C. Visaggi. 2014. Validation of taxon-specific sampling by novice collectors for studying drilling predation in fossil bivalves. Palaeogeography, Palaeoclimatology, Palaeoecology 412: 199-207.

Hausmann, I. M., H. Domanski, and M. Zuschin. 2018. Influence of setting, sieve size, and sediment depth on multivariate and univariate assemblage attributes of coral reef-associated mollusc death assemblages from the Gulf of Aqaba. Facies 64: 1-10.

Huntley, J. W. and M. Kowalewski. 2007. Strong coupling of predation intensity and diversity in the Phanerozoic fossil record: Proceedings of the National Academy of Sciences 104: 15006–15010.

Kelley, P. H., and T. A. Hansen. 1996. Naticid gastropod prey selectivity through time and the hypothesis of escalation. Palaios 11: 437-445.

Klompmaker, A. A., P. H. Kelley, D. Chattopadhyay, J. C. Clements, J. W. Huntley, and M. Kowalewski. 2019. Predation in the marine fossil record: studies, data, recognition, environmental factors, and behavior. Earth-science reviews 194: 472-520.

Kowalewski, M. 2002. The Fossil Record of Predation: An Overview of Analytical Methods. Pp. 3–42 *in* M. Kowalewski, P.H. Kelley, eds. The Fossil Record of Predation. Paleontological Society Papers, vol. 8. Yale University, New Haven.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Kowalewski, M., and A. P. Hoffmeister. 2003. Sieves and fossils: Effects of mesh size on paleontological patterns. *Palaios* 18: 460-469.

Langerhans, R. B. 2007. Evolutionary consequences of predation: avoidance, escape, reproduction, and diversification. Pp. 177-220 *in* Predation in organisms. Springer, Berlin, Heidelberg.

Nilsen, E. B., D. Christianson, J. M. Gaillard, D. Halley, J. DC. Linnell, M. Odden, M. Panzacchi, C. Toïgo, and B. Zimmermann. 2012. Describing food habits and predation: field methods and statistical considerations. Carnivore ecology and conservation: A handbook of techniques: 256-272.

Ottens, K. J., G. P. Dietl, P. H. Kelley, and S. D. Stanford. 2012. A comparison of analyses of drilling predation on fossil bivalves: bulk-vs. taxon-specific sampling and the role of collector experience. Palaeogeography, Palaeoclimatology, Palaeoecology 319:84-92.

Paine, R. T. 1966. Food web complexity and species diversity. The American Naturalist 100: 65-75.

Pringle, R. M., T. R. Kartzinel, T. M. Palmer, T. J. Thurman, K. Fox-Dobbs, C. CY Xu, M. C. Hutchinson, T.C. Coverdale, J.H. Daskin, D.A. Evangelista, and K.M. Gotanda. 2019. Predator-induced collapse of niche structure and species coexistence. Nature 570: 58-64.

Roy, K., D. J. Miller, and M. LaBarbera. 1994. Taphonomic bias in analyses of drilling predation: Effects of gastropod drill holes on bivalve shell strength. Palaios 9: 413-421.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Sarkar, D., G. Deole, S. Paul, S. Mondal, and R. Saha. 2021. Pre-burial taphonomic imprints on drilling intensity: a case study from the recent molluscs of Chandipur, India. Historical Biology 33: 1457-1463.

Signor, P. W., and Carlton E. B. 1984. The mid-Paleozoic precursor to the Mesozoic marine revolution. Paleobiology 10: 229-245.

Sime, J. A., and P. H. Kelley. 2016. Common mollusk genera indicate interactions with their predators were ecologically stable across the Plio-Pleistocene extinction. Palaeogeography, Palaeoclimatology, Palaeoecology 463: 216-229.

Smith, J. A., J. C. Handley, and G. P. Dietl. 2021. Accounting for uncertainty from zero inflation and overdispersion in paleoecological studies of predation using a hierarchical Bayesian framework. Paleobiology :1-18.

Stanley, S. M. 2008. Predation defeats competition on the seafloor. Paleobiology 34:1-21.

Vermeij, G. H. 1977. The Mesozoic marine revolution: evidence from snails, predators and grazers. Paleobiology 3: 245–258.

Vermeij, G. H. 1987. Evolution and Escalation: An Ecological History of Life. Princeton University Press, Princeton, New Jersey.

Visaggi, C. C., and P. H. Kelley. 2015. Equatorward increase in naticid gastropod drilling predation on infaunal bivalves from Brazil with paleontological implications. Palaeogeography, Palaeoclimatology, Palaeoecology 438: 285-299.

Zuschin, M., and R. J. Stanton Jr. 2001. Experimental measurement of shell strength and its taphonomic interpretation. Palaios 16: 161-170.

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table caption

Table 1. A summary of the model assemblages used for this study with varying evenness, predation intensity and predator preference.

Table 2. A summary of the published predation data from four Plio-Pleistocene fossil assemblages of Florida (Chattopadhyay and Baumiller, 2010).

Table 3. The result of Spearmann rank order correlation test for proportional abundance and PI_{prey} for the predation estimates across four Plio-Pleistocene fossil assemblages of Florida (Chattopadhyay and Baumiller, 2010). The statistically significant ($p<0.05$) results are marked in bold.

Table 4. The test-statistic (D) of Kolmogorov–Smirnov test comparing the predation estimates across four Plio-Pleistocene fossil assemblages of Florida using sample-standardization protocol. All the results are statistically significant ($p<0.05$).

Figure caption

Figure 1. An illustrative diagram of nine different molluscan communities with varying degrees of evenness (E_T) and predation intensity (PI_T). The predators attacking the prey are non-selective in their choice (Case 1). Mollusc drawings are from publicdomainpictures.net with subsequent modification. [Figure 1. Two-column; Grayscale]

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2. An illustrative diagram of nine different molluscan communities with varying degrees of evenness (E_T) and predation intensity (PI_T). The predators attacking the prey are selective and preferentially attack the most abundant prey species (Case 2). Mollusc drawings are from publicdomainpictures.net with subsequent modification. [Figure 2. Two-column; Grayscale]

Figure 3. An illustrative diagram of nine different molluscan communities with varying degrees of evenndegrees$_T$) and predation intensity (PI_T). The predators attacking the prey are selective in their choice and preferentially attack prey species rare in this assemblage (Case 3). Mollusc drawings are from publicdomainpictures.net with subsequent modification. [Figure 3. Two-column; Grayscale]

Figure 4. Plot showing variation in inferred predation intensity (PI_{inf}) with varying sample sizes for sizeserent model assemblages. The warmer colors represent higher evenness. The columns indicate increasing predation intensity in the original assemblage (PI_T) from left to right. The rows indicate the different degrees of the selectiveness of predation, denoted by Case 1, 2, and 3 from top to bottom. [Figure 4. Two columns; Color]

Figure 5. Plot showing deviation of the inferred predation intensity (PI_{inf}) from the overall predation intensity (PI_T) with varying sample size for the model assemblages representing Case 1 (non-selective predation). The rows indicate increasing evenness from top to bottom. The columns indicate increasing predation intensity in the original assemblage (PI_T) from left to right. The red line represents the overall predation intensity (PI_T), and the grey represents the simulated values for specific model assemblage. [Figure 4. Two columns; Color]

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 6. Plot showing deviation of the inferred predation intensity (PI_{inf}) from the overall predation intensity (PI_T) with varying sample size for the model assemblages representing Case 2 (selective predation of the common species). The rows indicate increasing evenness from top to bottom. The columns indicate increasing predation intensity in the original assemblage (PI_T) from left to right. The red line represents the overall predation intensity (PI_T), and the grey represents the simulated values for specific model assemblage. [Figure 6. Two columns; Color]

Figure 7. Plot showing deviation of the inferred predation intensity (PI_{inf}) from the overall predation intensity (PI_T) with varying sample size for the model assemblages representing Case 3 (selective predation of the rare species). The rows indicate increasing evenness from top to bottom. The columns indicate increasing predation intensity in the original assemblage (PI_T) from left to right. The red line represents the overall predation intensity ($PI_{T,}$) and the grey represents the simulated values for specific model assemblage. [Figure 7. Two columns; Color]

Figure 8. Plot showing variation in the inferred number of prey species ($S_{prey.inf}$) with varying sample sizes for different model assemblages. The warmer colors represent higher evenness. The columns indicate increasing predation intensity in the original assemblage (PI_T) from left to right. The rows indicate the different degrees of the selectiveness of predation, denoted by Case 1, 2, and 3 from top to bottom. [Figure 8. Two columns; Color]

Figure 9. Plot showing deviation (Dev_s) of the inferred prey species richness ($S_{prey.inf}$) from the original prey species richness (S_{prey}) with varying sample size for the model

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

assemblages representing Case 1 (non-selective predation). The rows indicate increasing evenness from top to bottom. The columns indicate increasing predation intensity in the original assemblage (PI_T) from left to right. The red line represents the original prey species richness (S_{prey}), and the grey represents the simulated values for specific model assemblage. [Figure 9. Two columns; Color]

Figure 10. Plot showing deviation (Dev_s) of the inferred prey species richness ($S_{prey.inf}$) from the original prey species richness (S_{prey}) with varying sample size for the model assemblages representing Case 2 (selective predation of the common species). The rows indicate increasing evenness from top to bottom. The columns indicate increasing predation intensity in the original assemblage (PI_T) from left to right. The red line represents the original prey species richness (S_{prey}), and the grey represents the simulated values for specific model assemblage. [Figure 10. Two columns; Color]

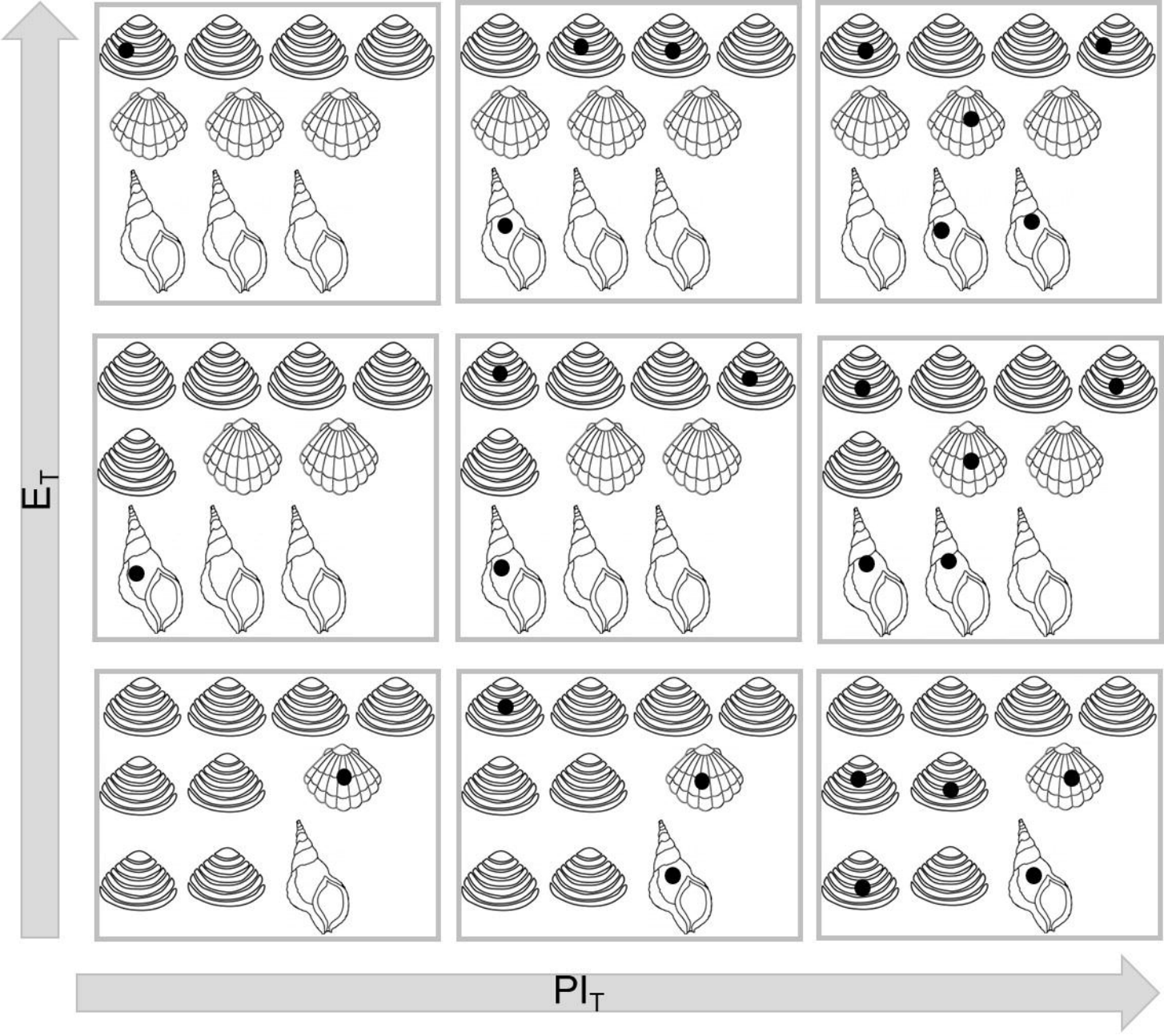
Figure 11. Plot showing deviation (Dev_s) of the inferred prey species richness ($S_{prey.inf}$) from the original prey species richness (S_{prey}) with varying sample size for the model assemblages representing Case 3 (selective predation of the rare species). The rows indicate increasing evenness from top to bottom. The columns indicate increasing predation intensity in the original assemblage (PI_T) from left to right. The red line represents the original prey species richness (S_{prey}), and the grey represents the simulated values for specific model assemblage. [Figure 11. Two columns; Color]

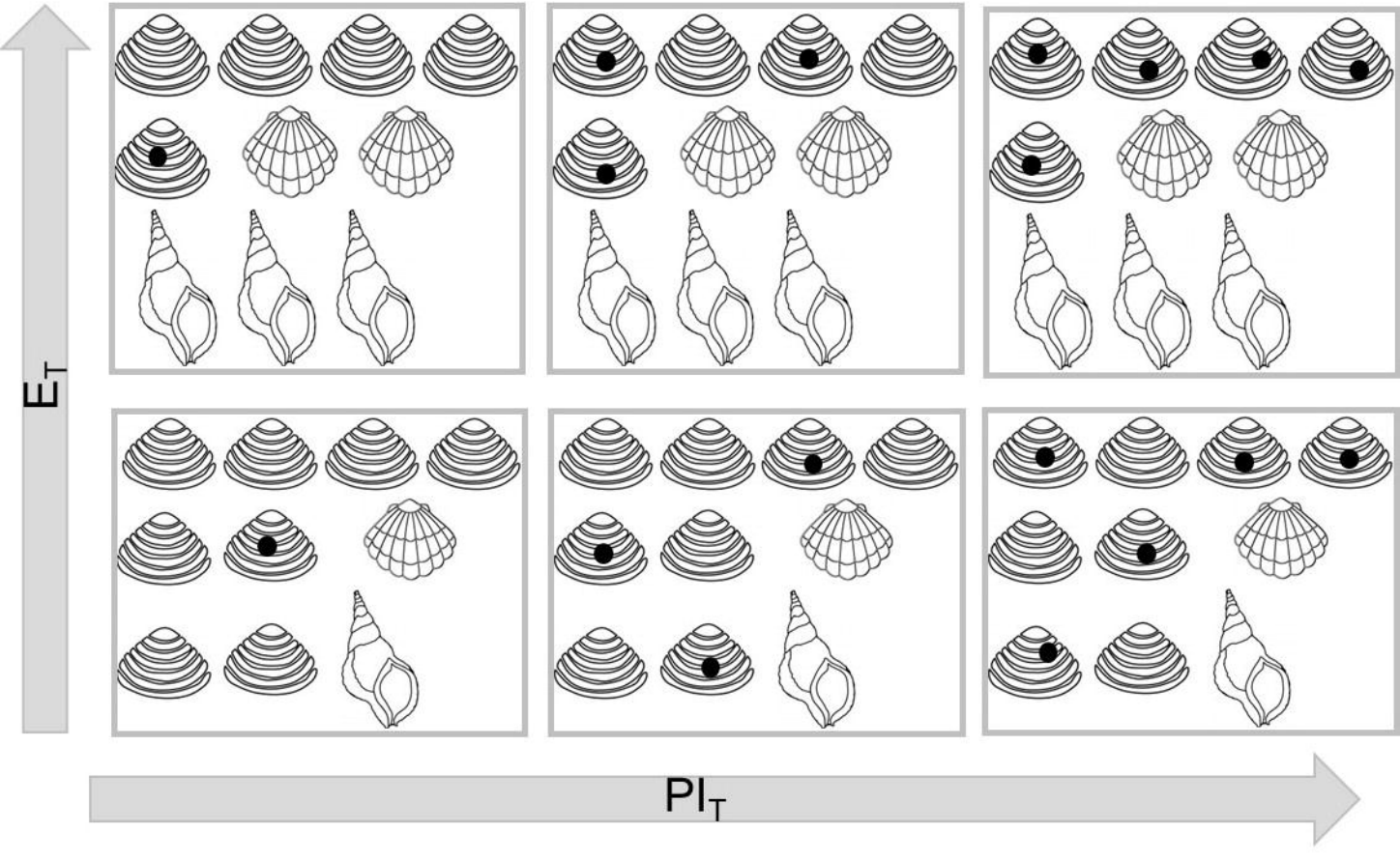
Figure 12. The plot shows variation in inferred predation intensity (PI_{inf}) and inferred the number of prey species ($S_{prey.inf}$) with specific sample sizes for different model

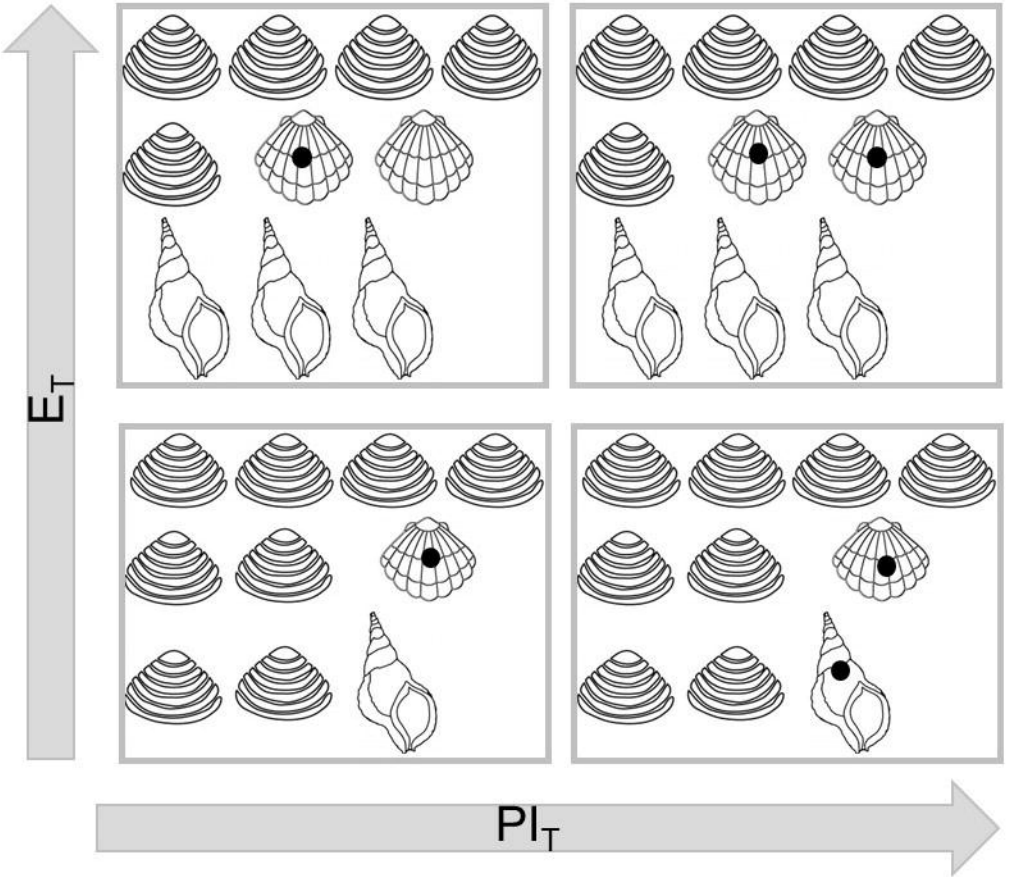
bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

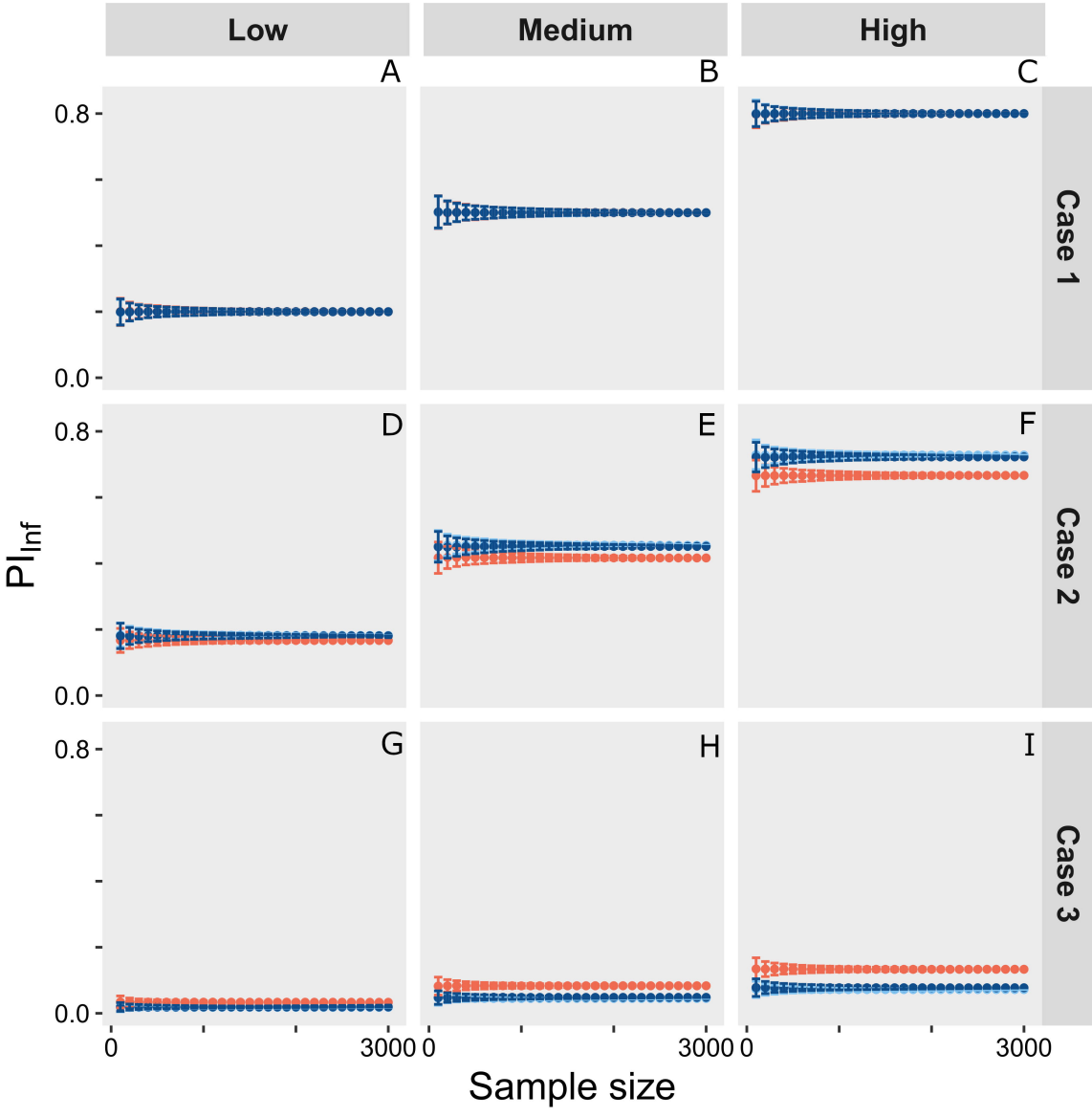
assemblages. The warmer colors represent higher evenness. The columns indicate increasing predation intensity in the original assemblage (PI_T) from left to right. The rows indicate the different degrees of the selectiveness of predation, denoted by Case 1, 2, and 3 from top to bottom. [Figure 12. Two columns; Color]

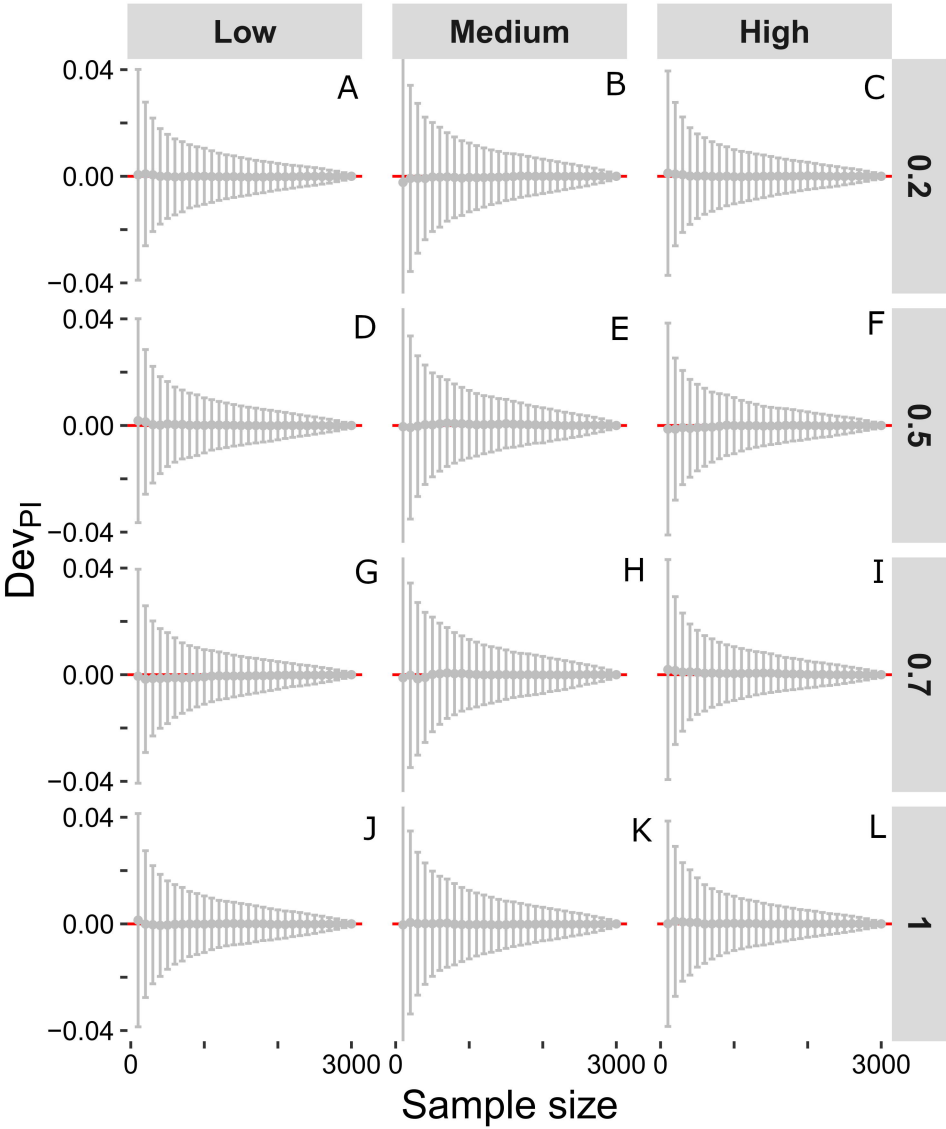
Figure 13. Plot showing variation in inferred estimates of drilling and durophagous predation with varying degrees of sampling for four Pleistocene molluscan assemblages of Florida with different evenness (E_T). The top row represents the sample size variation in inferred predation intensity (PI_{inf}). The middle row shows the inferred number of prey species ($S_{prey.inf}$) with varying sample sizes. The bottom row represented the relationship between the inferred predation intensity (PI_{inf}) and inferred number of prey species ($S_{prey.inf}$) at specific sizes for the four assemblages. [Figure 13. Two columns; Color]

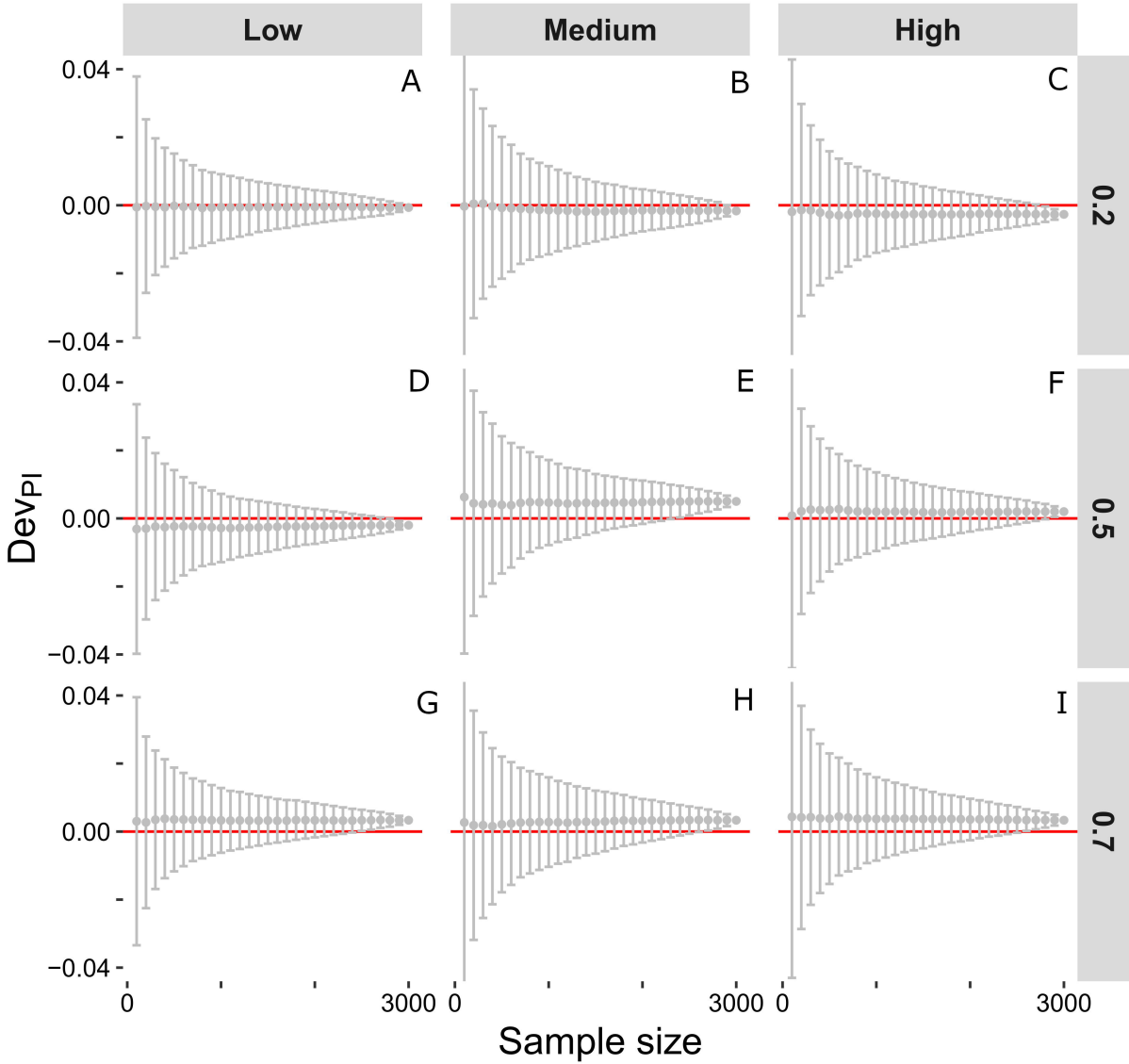

E_T
PI_T

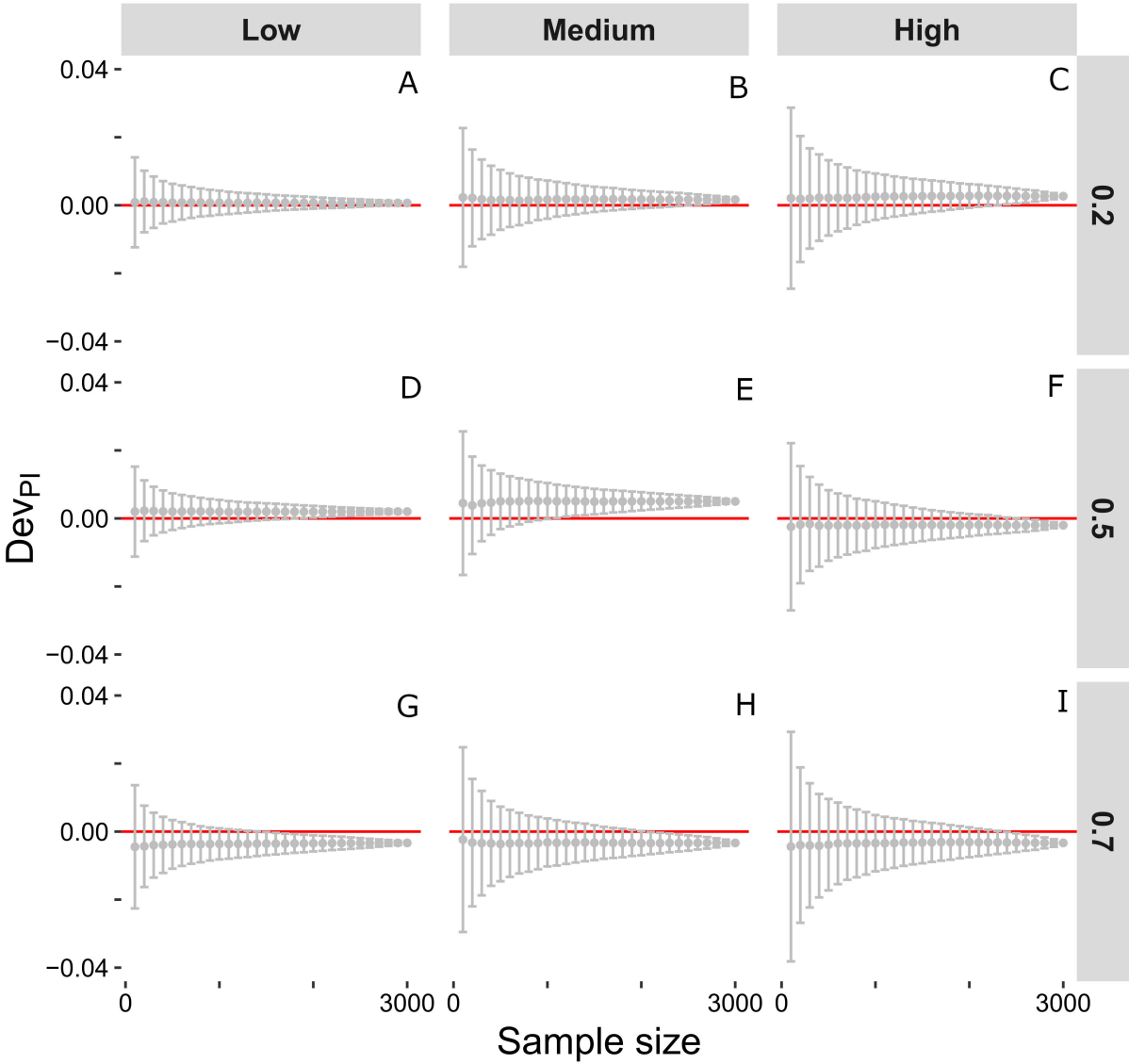

E_T
PI_T

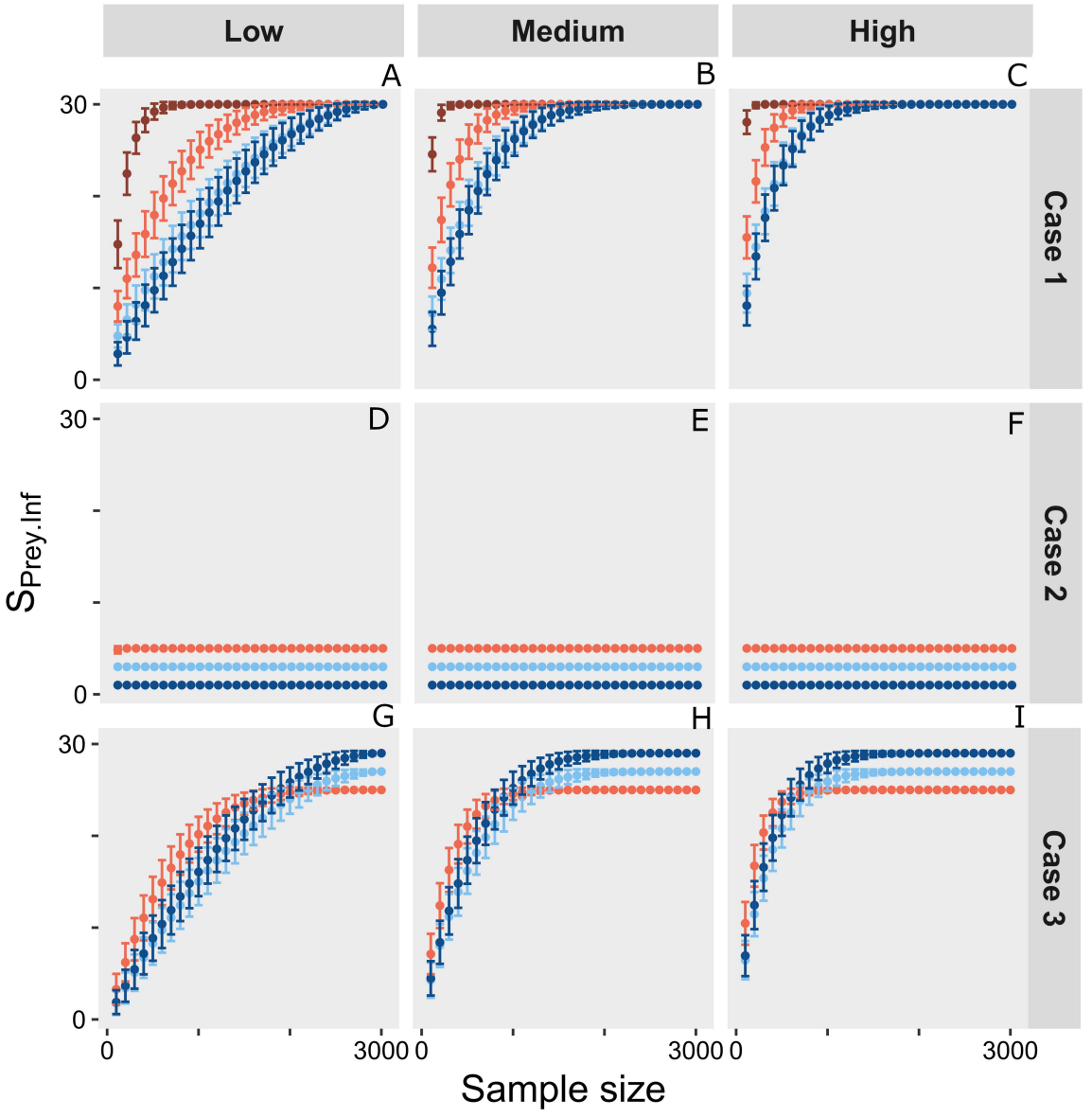

E_T
PI_T

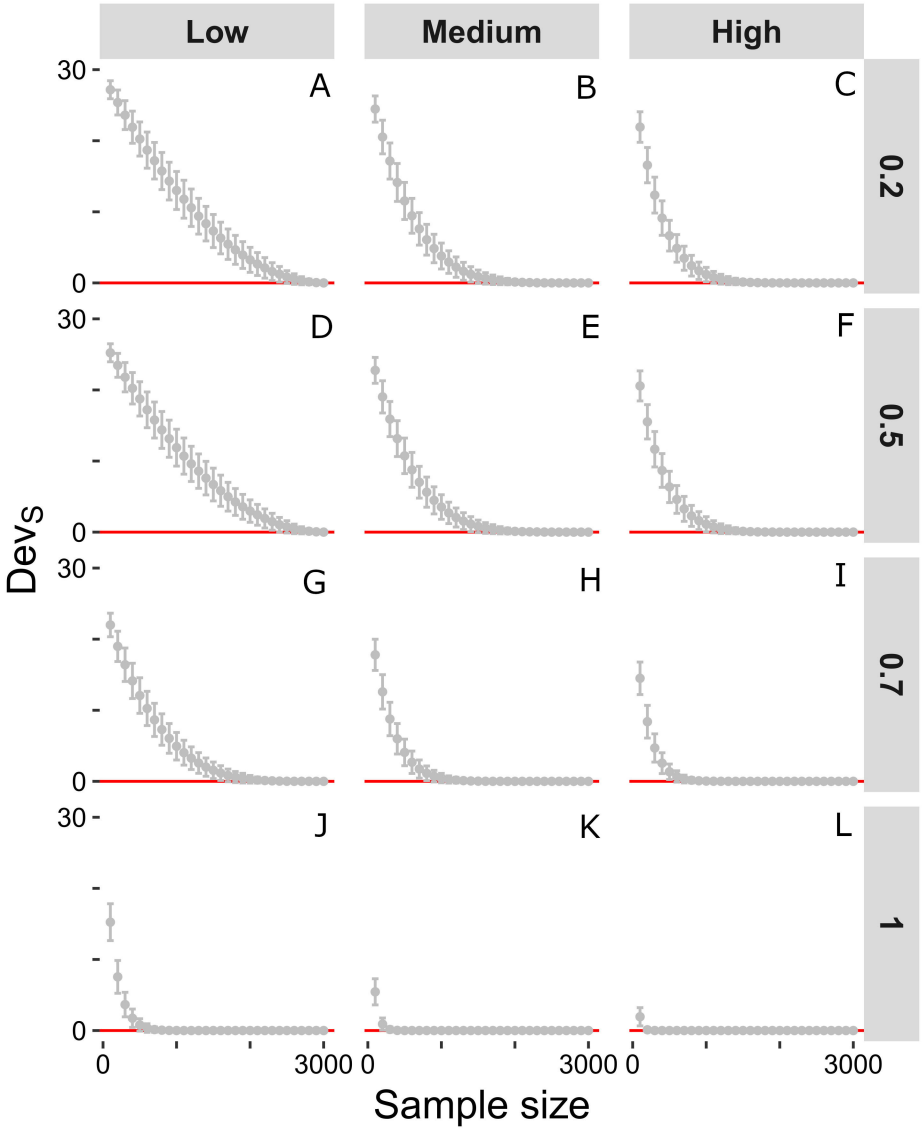

Low
Medium
High
A
B
C
D
E
F
G
H
I
Case 1
Case 2
Case 3
0.8
0.0
PI_{Inf}
0
3000
Sample size

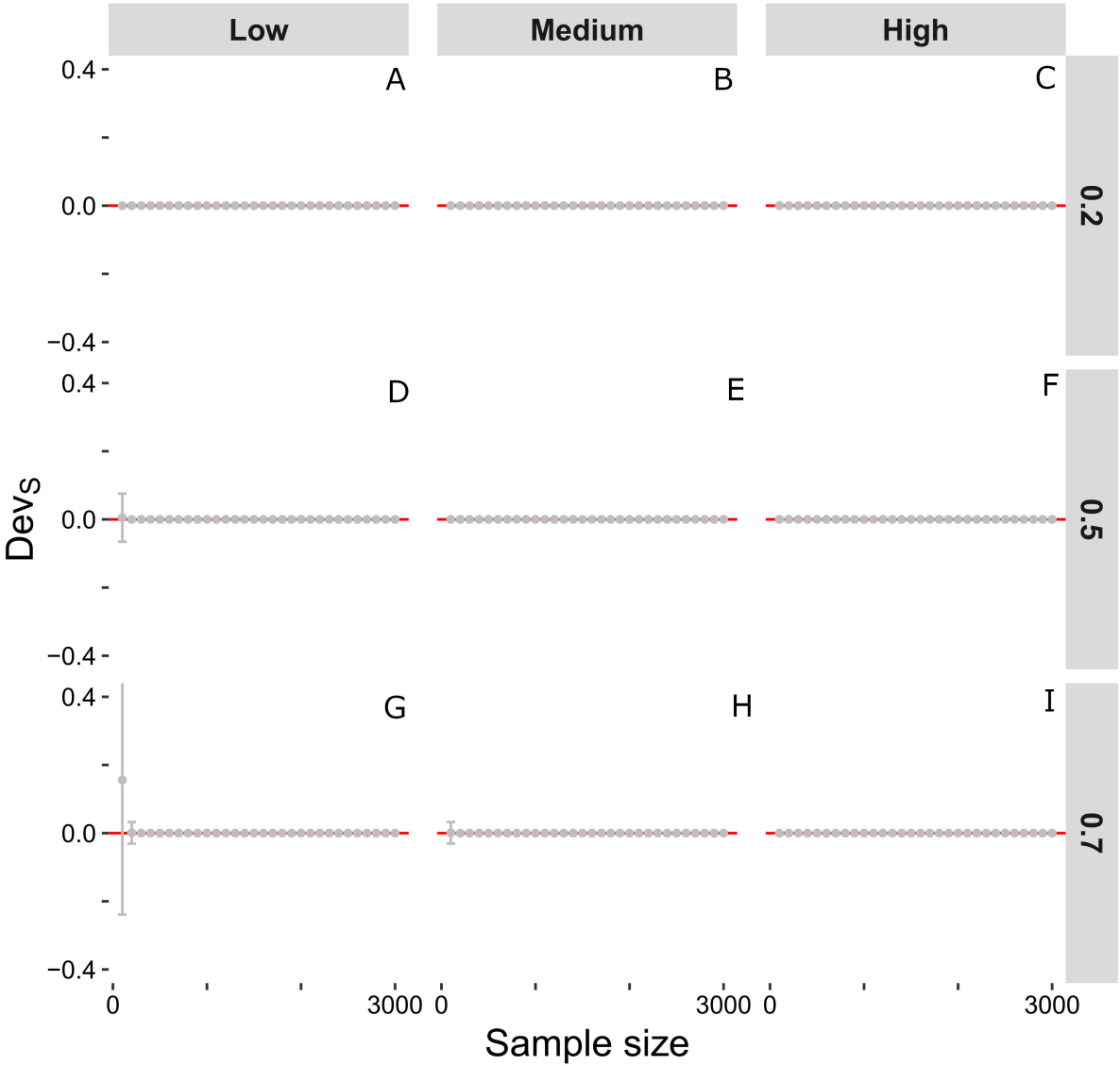

Low
Medium
High
0.2
0.5
0.7
1
A
B
C
D
E
F
G
H
I
J
K
L
0.04
0.00
−0.04
0
3000
Dev_PI
Sample size

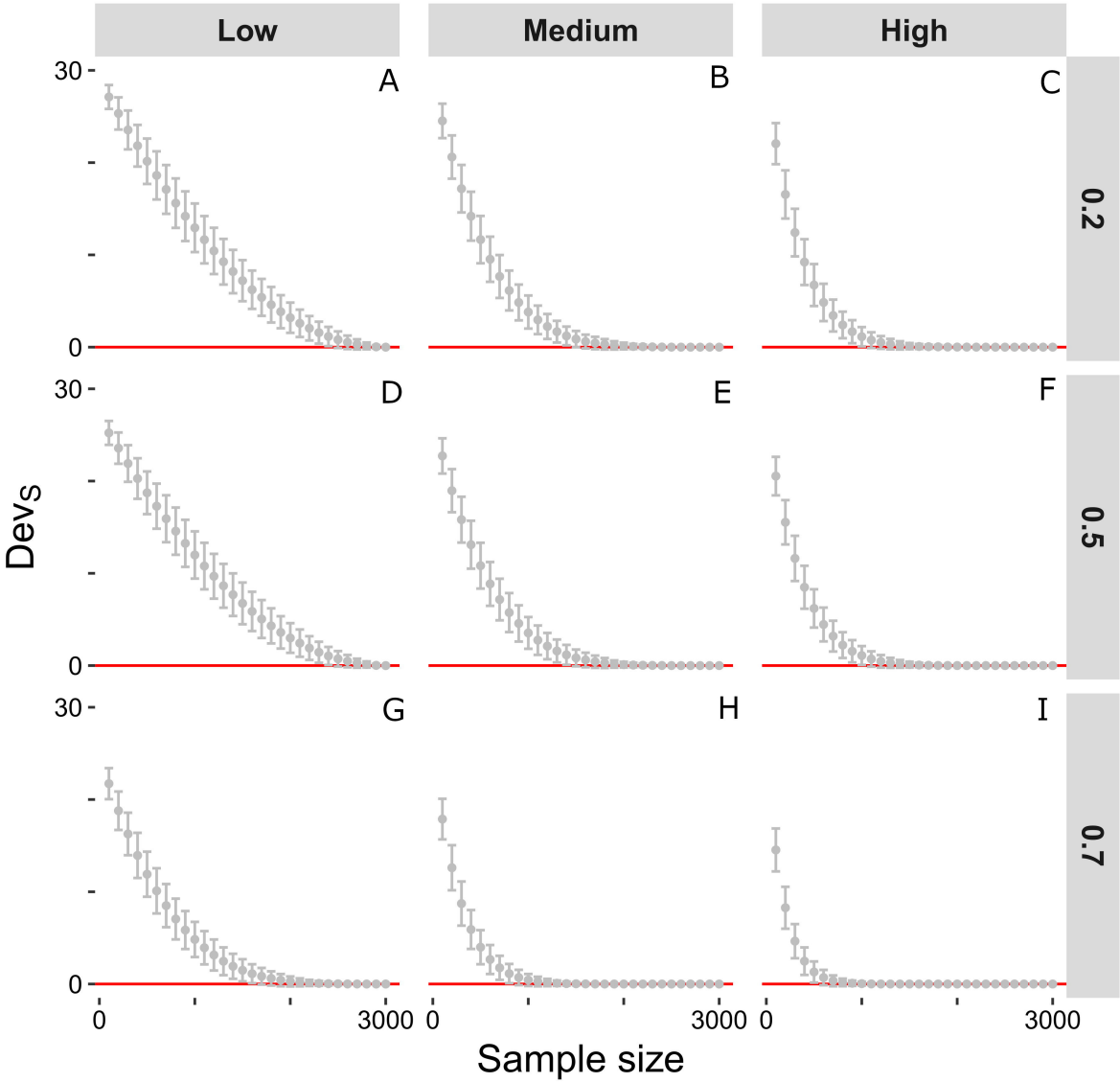

Low
Medium
High
A
B
C
D
E
F
G
H
I
0.2
0.5
0.7
0.04
0.00
−0.04
Dev_{PI}
0
3000
Sample size

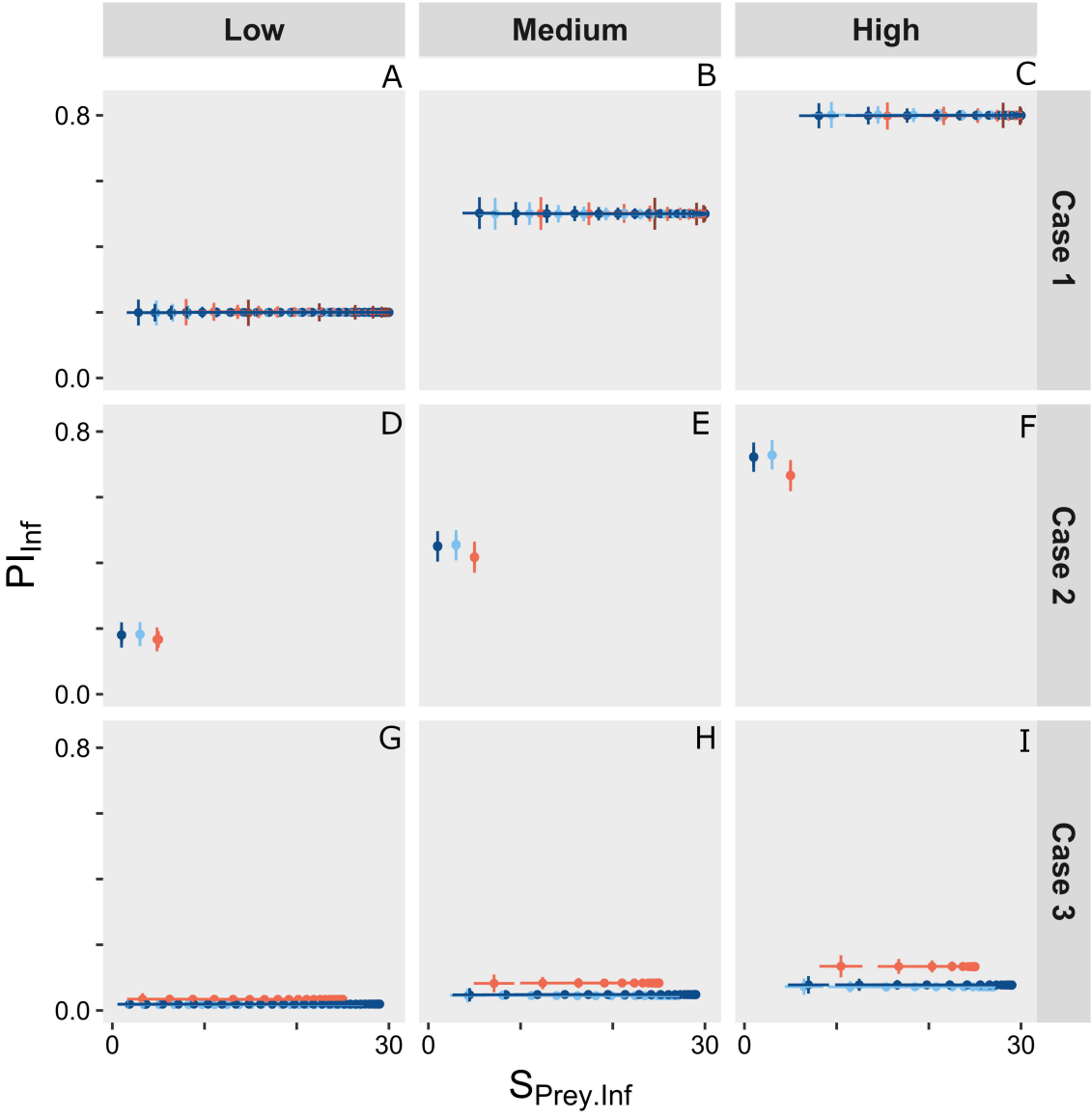

Low
Medium
High
0.2
0.5
0.7
A
B
C
D
E
F
G
H
I
0.04
0.00
−0.04
Dev_{PI}
0
3000
Sample size


Low
Medium
High
A
B
C
D
E
F
G
H
I
Case 1
Case 2
Case 3
30
0
$S_{Prey.Inf}$
0
3000
Sample size


Low
Medium
High
0.2
0.5
0.7
1
A
B
C
D
E
F
G
H
I
J
K
L
30
0
3000
Dev$_S$
Sample size


Low
Medium
High
A
B
C
D
E
F
G
H
I
0.2
0.5
0.7
0.4
0.0
−0.4
Dev_S
0
3000
Sample size


Low
Medium
High
0.2
0.5
0.7
A
B
C
D
E
F
G
H
I
30
0
3000
Dev$_S$
Sample size


Low
Medium
High
A
B
C
D
E
F
G
H
I
Case 1
Case 2
Case 3
0.8
0.0
0
30
PI_{Inf}
$S_{Prey.Inf}$

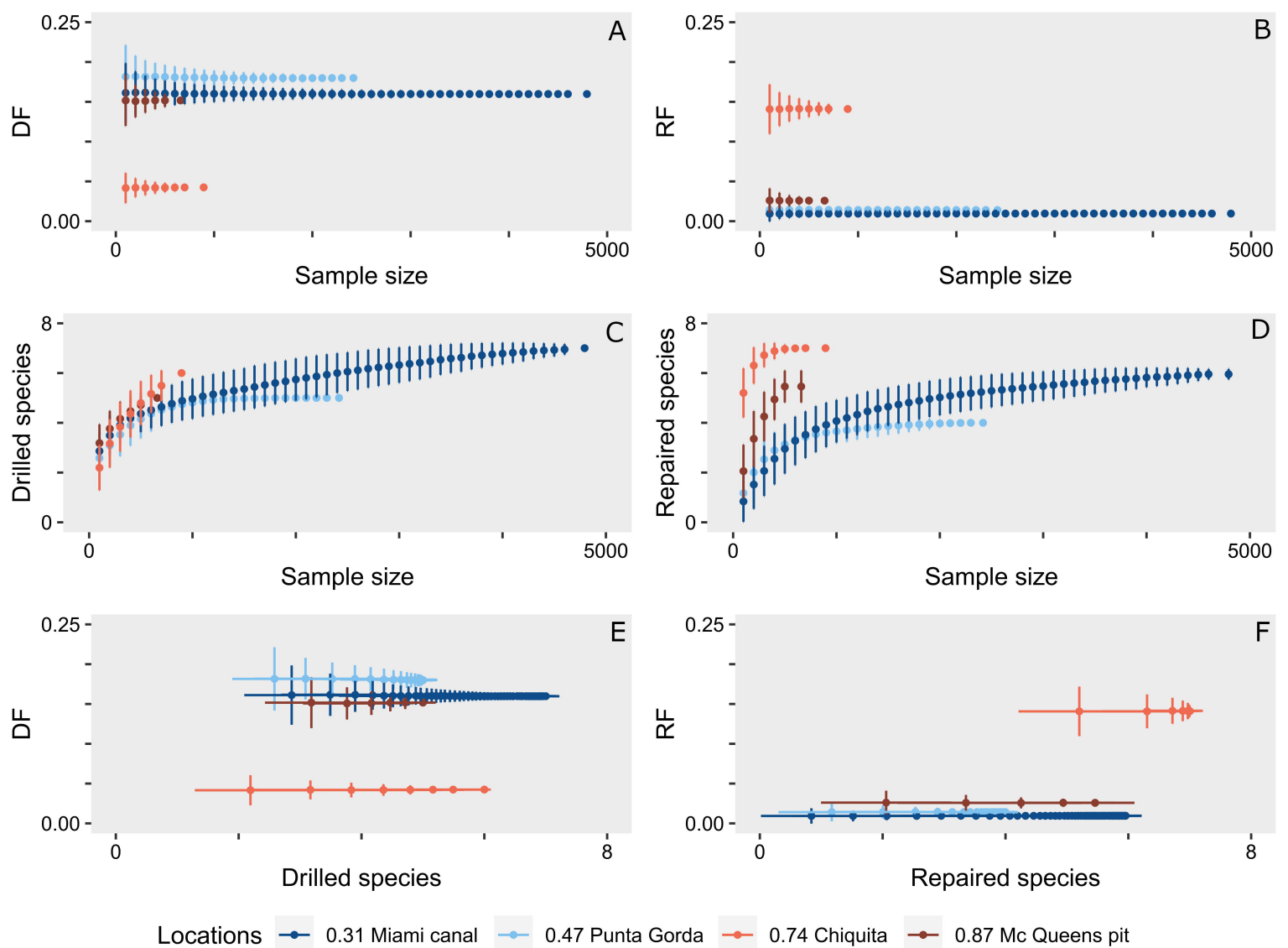

A
B
C
D
E
F
DF
RF
Drilled species
Repaired species
Sample size
0.25
0.00
0
5000
8
Locations
0.31 Miami canal
0.47 Punta Gorda
0.74 Chiquita
0.87 Mc Queens pit

Table 1. A summary of the model assemblages used for this study with varying evenness, predation intensity and predator preference.

Evenness	Structure of the model community	Case 1 (Preys of all species are attacked equal probablity)				Case 2 (Preys of only common species are attacked)				Case 3 (Preys of only rare species are attacked)			
		Low (PI_{prey}=0.2)	Medium (PI_{prey}=0.5)	High (PI_{prey}=0.8)	S_{prey}	Low (PI_{prey}=0.2)	Medium (PI_{prey}=0.5)	High (PI_{prey}=0.8)	S_{prey}	Low (PI_{prey}=0.2)	Medium (PI_{prey}=0.5)	High (PI_{prey}=0.8)	S_{prey}
E_T = 0.2	$N_{(S=1:29)}$=10, $N_{(S=30)}$=2710 [1*2710 + 29*10] = 3000	PI_T=0.2	PI_T=0.5	PI_T=0.8	30	PI_T=0.18	PI_T=0.45	PI_T=0.72	1	PI_T=0.02	PI_T=0.05	PI_T=0.08	29
E_T = 0.5	$N_{(S=1:3)}$=910, $N_{(S=4:30)}$=10 [3*910 + 27*10] = 3000	PI_T=0.2	PI_T=0.5	PI_T=0.8	30	PI_T=0.18	PI_T=0.46	PI_T=0.73	3	PI_T=0.02	PI_T=0.05	PI_T=0.07	27
E_T = 0.7	$N_{(S=1:5)}$=500, $N_{(S=6:30)}$=20 [5*500 + 25*20] = 3000	PI_T=0.2	PI_T=0.5	PI_T=0.8	30	PI_T=0.17	PI_T=0.42	PI_T=0.67	5	PI_T=0.03	PI_T=0.08	PI_T=0.13	25
E_T = 1	$N_{(S=1:30)}$=100 [30*100] = 3000	PI_T=0.2	PI_T=0.5	PI_T=0.8	30	NA	NA	NA	NA	NA	NA	NA	NA

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. A summary of the published predation data from four Plio-Pleistocene fossil assemblages of Florida (Chattopadhyay and Baumiller, 2010).

Locality	Evenness (E_T)	Sample size	S_T	Drilling frequency	$S_{prey.drill}$	Repair scar frequency	$S_{prey.repair}$
Miami Canal	0.31	4794	7	0.16	7	0.01	6
Punta Gorda	0.47	2417	5	0.18	5	0.01	4
Chiquita	0.74	894	7	0.04	6	0.14	7
McQueen's pit	0.87	657	6	0.15	5	0.03	6

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 3. The result of Spearmann rank order correlation test for proportional abundance and PI_{prey} for the predation estimates across four Plio-Pleistocene fossil assemblages of Florida (Chattopadhyay and Baumiller, 2010). The significant results are marked in bold.

Predation	Location	rho	p	Inferred scenario
Drilling	Punta Gorda	0.87	0.05	Case 2
	McQueen's pit	0.83	0.06	Case 1
	Chiquita	0.68	0.08	Case 1
	Miami canal	**0.99**	**<0.001**	Case 2
Durophagy	Punta Gorda	0.21	0.74	Case 1
	McQueen's pit	0.46	0.35	Case 1
	Chiquita	0.24	0.61	Case 1
	Miami canal	**0.79**	**0.03**	Case 2

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 4. The test-statistic (D) of Kolmogorov–Smirnov test comparing the predation estimates across four Plio-Pleistocene fossil assemblages of Florida using sample-standardization protocol. All the results are statistically significant ($p<0.05$).

Estimate	Predation	Location	McQueen's pit	Chiquita	Miami Canal
Predation intensity	Drilling	Punta Gorda	**0.8**	**1**	**0.53**
		McQueen's pit		**0.24**	**0.31**
		Chiquita			**1**
	Durophagy	Punta Gorda	**0.9**	**1**	**0.19**
		McQueen's pit		**0.96**	**0.85**
		Chiquita			**1**
Prey species richness	Drilling	Punta Gorda	**0.38**	**0.31**	**0.13**
		McQueen's pit		**0.24**	**0.36**
		Chiquita			**0.29**
	Durophagy	Punta Gorda	**0.84**	**1**	**0.41**
		McQueen's pit		**1**	**0.95**
		Chiquita			**1**

bioRxiv preprint doi: https://doi.org/10.1101/2022.07.18.500550; this version posted July 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.