Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Taxonomically standardized iNaturalist records reveal broad-scale patterns in documented polychaete diversity across U.S. coastlines

Andrew Davinack

davinack_drew@wheatoncollege.edu

Wheaton College Massachusetts

Leslie Harris

Natural History Museum of Los Angeles County

Research Article

Keywords: community, science, annelida, WoRMS, marine, biogeography

Posted Date: August 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10619974/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Citizen-science platforms are increasingly used to document biodiversity, but their value for taxonomically difficult marine invertebrates remains uncertain. Polychaete annelids are diverse, ecologically important, and often difficult to identify from photographs alone, raising questions about how community-science records should be interpreted. Here, we evaluated research-grade iNaturalist observations of Polychaeta across the continental United States. We retrieved 13,775 observations from iNaturalist, standardized names against a World Register of Marine Species-derived taxonomic backbone, assigned records to coastal regions using geographic coordinates, and analyzed observation-based richness, family composition, species overlap, rank-observation structure, and rarefaction patterns. After filtering, 12,539 observations representing 302 species, 54 families, and 4,027 observers were retained. Sampling effort was highly uneven, with the West Coast accounting for most observations, species, families, and observers. Species overlap among coasts was limited, and dominant families differed regionally, although tubiculous families such as the Serpulidae were prominent across all three coasts. Rank-observation curves showed that a small number of frequently documented species dominated each regional record, whereas many taxa were represented by few observations. Observation-based rarefaction indicated continued species accumulation, especially on the West Coast. These results show that taxonomically standardized iNaturalist records can reveal broad patterns in documented polychaete diversity while underscoring the need for expert validation, voucher-based taxonomy, and careful interpretation of observation-derived biodiversity data.

Introduction

Marine biodiversity is declining rapidly under intensifying anthropogenic pressures, including climate change, habitat degradation, pollution, and the spread of invasive species (IPBES, 2019; Ward et al., 2022; Jaureguiberry et al., 2022). Yet this loss is occurring against a backdrop of profound uncertainty: a substantial proportion of marine life remains undescribed, and the true scale of ocean biodiversity is still unknown (Mora et al., 2011; Costello & Chaudhary, 2017). Many marine taxa harbor extensive cryptic diversity – genetically distinct lineages that remain masked under conserved morphologies, with estimates suggesting that tens of thousands of marine species await formal discovery (Nygren, 2014; Struck & Cerca, 2019; Cahill et al., 2024). The consequence is a paradoxical situation in which marine biodiversity is disappearing faster than it can be documented (Hending, 2024).

This crisis is compounded by the *taxonomic impediment*: the global shortage of appropriately trained taxonomists and systematists, particularly for species-rich and morphologically challenging groups, that constrains the discovery, delimitation, and reliable identification of species (Engel et al., 2021). Declines in funding, shrinking expertise, and the loss of institutional support have disproportionately affected “problematic” taxa such as annelids, nematodes, and many micro-/meiofaunal lineages (Paknia et al., 2015). As a result, some authors have proposed alternative approaches to accelerate the description of diversity, including genetic lineage cataloging, automated delimitation pipelines, and taxonomic shortcuts that relax evidentiary standards such as type designation (Ranasinghe et al., 2022; Vivien et al.,

2025; Davinack & Fitzhugh, 2025). While these approaches highlight the urgency of the problem, none supersede the foundational role of biodiversity surveys – systematic inventories of species occurrences that underpin biogeography, conservation, and ecological understanding. Rapid assessment methods such as bioblitzes, coastal biodiversity inventories, and regional monitoring programs continue to be central to documenting species presence and tracking distributional changes across space and time (Pederson et al. 2021; Davinack et al., 2024; Fowler et al., 2025).

In recent years, citizen science – the participation of members of the public in the collection, documentation, or analysis of scientific data – has emerged as a powerful supplement to traditional biodiversity survey efforts (Dickinson et al., 2012; Theobald et al., 2015; Chandler et al., 2017). Among digital citizen-science platforms, iNaturalist has become one of the world's largest repositories of georeferenced species observations (Mason et al, 2025). Users upload photographs and metadata, and identifications are refined through a large expert community. When multiple users agree on an identification and data quality thresholds are met, observations achieve 'research grade' status and are integrated into global biodiversity portals such as GBIF. Recently, iNaturalist data have been incorporated into studies of species richness (Marsh et al., 2026), invasive species spread and management (Potgieter et al., 2024; Cull, 2025), range shifts (Wagner et al., 2025), phenology (Dinnage et al., 2025), plant-pollinator interactions (Marin-Gomez et al., 2022), and assessments of taxon-specific diversity patterns (Rosa et al., 2022; Herrera et al., 2025). Despite limitations, including uneven sampling effort, observer bias, geographic accessibility, and variable detectability (Szabolcs et al., 2026) – iNaturalist has proven to be a valuable source for biodiversity assessments when data are interpreted with appropriate caution.

Polychaetes represent one of the most ecologically and taxonomically diverse lineages in the marine realm, with more than 20,000 described species, and many more awaiting discovery (Rouse et al., 2022). They occupy nearly every marine habitat, from intertidal zones to the deep sea, and include numerous cryptic species complexes that remain unresolved even in well-studied regions (Nygren, 2014; Rouse et al., 2022). Their morphological diversity, ecological breadth, and propensity for invasions make them both biologically and ecologically important but also taxonomically challenging (Nygren, 2014). Yet, despite their ubiquity, polychaetes remain underrepresented in citizen-science literature, and as far as we know, no comprehensive assessment of polychaete records in iNaturalist has been conducted thus far.

For polychaetes, the value of community-science records depends not only on the number of observations but also on the extent to which those records can be reconciled with authoritative taxonomic frameworks. Integrating iNaturalist observations with curated taxonomic backbones such as the World Register of Marine Species (WoRMS) provides a way to evaluate documented diversity while maintaining a clear distinction between observation-derived records and formal taxonomic evidence.

Here, we evaluated the structure and diversity of polychaete observations from iNaturalist across the three major U.S. coastlines: the East Coast (western Atlantic), Gulf Coast, and West Coast (Pacific). Specifically, we (1) quantify observed species richness and taxonomic composition across regions, (2)

assess the structure of observation frequencies using rank-observation distributions, (3) examine sampling completeness, and (4) evaluate biogeographic overlap via shared-species metrics. By characterizing patterns of observation-derived diversity in one of the ocean's most speciose invertebrate groups, this study provides the first baseline assessment of polychaete diversity as captured through citizen science and highlights both the potential and the limitations of using iNaturalist for marine biodiversity monitoring.

Materials and Methods

Data acquisition from iNaturalist

A comprehensive dataset of research-grade (Annelida: Polychaeta) observations from the iNaturalist platform (www.inaturalist.org), was compiled from 2008–2025. Observations were retrieved using the Python package *pyinaturalist* (v.0.19) by querying the taxon Polychaeta (iNaturalist taxon ID: 47490). Only records classified as "research grade" – requiring georeferenced photographic evidence, a community-confirmed taxon ID, and complete metadata – were included to ensure consistent data quality. Geographic filtering was implemented by retaining only records with coordinates falling within the bounding box of the continental United States and its coastal margins (latitudes 24–49°N; longitudes –125 to – 66°W). This ensured that all records included in downstream analyses were drawn from the U.S. East Coast, Gulf Coast, or West Coast biogeographic regions. For each observation, the following metadata were extracted: observation ID, observation timestamp, latitude, longitude, place description, observer username, iNaturalist taxon ID, scientific name, common name (if available), and assigned taxonomic rank. A total of 13,775 U.S. research-grade polychaete observations were retrieved prior to taxonomic cleaning and spatial classification.

WoRMS taxonomic standardization

To ensure authoritative and internally consistent taxonomic assignments across all observations, we standardized every scientific name in the dataset against the World Register of Marine Species (WoRMS). We used a curated WoRMS taxonomic backbone (obtained from the Flanders Marine Institute) that included accepted names, synonymies, authorship, AphiaIDs, and complete higher-level classifications (see data availability link). Because iNaturalist identifications may include outdated nomenclature, informal names, or infraspecific epithets, we implemented a multi-step reconciliation procedure. Scientific names were first matched directly against the 'ValidName' field in WoRMS, and infraspecfic taxa (e.g. *Spiochaetopterus costarum pottsi*) were reduced to binomial form to improve alignment. Records identified only to genus were assigned family-level classification using a genus-family lookup table derived from the WoRMS hierarchy. Subfamilial identifiers, such as Spirorbinae or Nereidinae, were mapped to their corresponding families to ensure consistency with current taxonomic usage. A small number of historically problematic or ambiguous names – such as *Spirorbis spirorbis*, which has undergone multiple revisions – were reviewed and corrected manually following WoRMS

taxonomic conventions. After reconciliation, 13,774 of the 13,775 observations (99.99%) were successfully linked to an authoritative family-level name, with only one record remaining unresolved due to insufficient taxonomic information.

Data cleaning and preparation

Observations lacking geographic coordinates or containing malformed metadata were removed prior to analysis. All remaining fields were cleaned and standardized for consistency, including normalization of scientific name formatting and generation of a canonical binomial name for cross-referencing. The final dataset retained, for every observation, a complete suite of metadata including observation ID, datetime, geographic coordinates, observer username, taxon ID, scientific name, matched binomial, and WoRMS-validated higher-level taxonomy (genus through kingdom).

Spatial classification into U.S. coastal regions

Because iNaturalist place-name metadata are heterogeneous and often inconsistently formatted, spatial classification was performed using geographic coordinates rather than textual locality fields. Observations were assigned to one of three U.S. coastal regions – East Coast, Gulf Coast, and West Coast – based on their latitude and longitude. West Coast observations were defined as those occurring west of -100° longitude. Gulf Coast records were restricted to observations between − 100° and − 75° longitude and south of 31°N latitude, capturing the marine shorelines of Texas, Louisiana, Mississippi, Alabama and the Florida Panhandle. East Coast observations were defined as all records occurring east of -75° longitude. This coordinate-based approach ensures a reproducible spatial classification that aligns with major U.S. marine biogeographic boundaries (Fig. 1).

Diversity and community structure analyses

Observed diversity patterns were quantified independently for each coastal region using species-level observation counts derived from the curated dataset. For every coast, we calculated total species richness, family richness, and the number of contributing observers to characterize sampling breadth. To assess patterns of observed diversity, we computed Shannon diversity (H’), Simpson diversity (1-D), and Pielou’s evenness (J') using species level-abundances. Because iNaturalist observations represent opportunistic sampling rather than systematic survey effort, these metrics reflect observation-based diversity rather than true ecological community structure, but they nonetheless provide comparative insight into how taxonomic representation differs across regions.

The structure of observation frequencies was further examined using rank-observation distributions. For each coast, species were ranked according to their number of iNaturalist observations, and observation frequencies were plotted on a logarithmic scale. These curves allowed us to compare how strongly each regional observation record is shaped by a few highly photographed taxa versus more evenly

represented communities. Differences in curve steepness between coasts provided a qualitative measure of taxonomic dominance and sampling heterogeneity.

To evaluate whether observed species-level richness approached saturation within the iNaturalist observation record, we generated observation-based rarefaction curves for each species. For each sampling depth, observation records were repeatedly subsampled without replacement from the species-by-observation dataset, using 100 randomization replicates per depth. Mean expected richness and standard deviation were then calculated across replicates. Complementary species accumulation curves were created to illustrate the rate at which additional species were documented as the cumulative number of observations increased. Together, these analyses distinguished well-sampled assemblages from those where additional sampling would likely yield substantial numbers of newly recorded taxa.

Finally, to quantify biogeographic turnover among coastal regions, we compared species lists across coasts and calculated the number of species shared among all three regions, shared by exactly two regions or unique to a single coast. These overlap patterns provide a broad-scale assessment of regional distinctiveness and reveal the extent to which polychaete taxa observed on iNaturalist exhibit restricted versus widespread distributions across U.S. marine coastlines.

Results

A total of 13,775 research-grade iNaturalist observations of polychaete worms were initially retrieved. After excluding records lacking geographic coordinates or falling outside the three focal U.S. coastal regions, 12,539 observations were retained for analysis. These represented 302 species across 54 families, contributed by 4,027 observers. Sampling effort varied considerably among coasts. The West Coast accounted for the majority of observations (10,906 records, 198 species, 36 families, 3,224 observers), and identification activity in this region was dominated by a single expert identifier (*leslieh*; author L.H.), who contributed 810 polychaete identifications. The East Coast contributed 780 observations (61 species, 28 families, 460 observers) while the Gulf Coast contributed 853 observations (43 species, 18 families, 343 observers).

Species overlap among coasts was limited. Only seven species were recorded on all three coasts. These were *Alitta succinea*, *Arenicola cristata, Arenicola marina*, *Ficopomatus enigmaticus, Hydroides elegans* and *Myrianida pachycera*. When genus-level records were included, *Diopatras* sp. was also shared among all three coasts. This represents the strict intersection of East, Gulf and West Coast species lists. Pairwise comparisons showed greater overlap, with the East and Gulf Coasts sharing more species with one another than with the west coast. Of these, four taxa were not identified to the species level (*Alitta* sp., *Diopatra* sp., *Myxicola* sp., and Spirorbinae sp.). Observation frequencies of shared species also varied substantially among regions, with several taxa showing pronounced coast-specific dominance (Fig. 2).

Despite these modest similarities, each coast was characterized by a distinct taxonomic composition. The most frequently observed families varied among coasts with the West Coast exhibiting the greatest family-level richness and diversity of dominant families. For the West Coast, the most frequently observed families included Sabellariidae, Serpulidae, and Sabellidae, for the East Coast they were Nereididae, Serpulidae, and Arenicolidae, and for the Gulf Coast, they were Amphinomidae, Serpulidae, and Sabellidae (Fig. 3).

At the species-level, the most frequently observed species were *Phragmatopoma californica*, *Serpula columbiana* and *Eudistylia vancouveri* for the West Coast, *Hermodice carunculata*, *Spirobranchus giganteus*, and *Chloeia euglochis* for the Gulf Coast, and *Alitta succinea, Spirorbis spirorbis* and *Arenicola cristata* for the East Coast (Fig. 4).

Rank-observation curves revealed that all three coastal records were dominated by a small number of frequently documented taxa (Fig. 5). When plotted on a logarithmic scale, the West Coast curve extended to nearly 200 ranked species, displaying a long tail composed of numerous taxa represented by few observations, as well as several highly photographed species. The East Coast and Gulf Coast curves showed steeper declines after the most frequently observed taxa and shorter tails largely composed of singletons and doubletons, reflecting the smaller number of species recorded in these regions. All three curves showed consistent ordering of relative abundance within each coast.

Diversity indices also varied among coasts. Species richness was highest on the West Coast (S = 198), intermediate on the East Coast (S = 61), and lowest on the Gulf Coast (S = 43). Shannon diversity values were 3.05, 2.70, and 2.59 for the West, East, and Gulf Coasts, respectively. Simpson diversity values were similar across regions (1-D = 0.883–0.887). Pielou's evenness differed more strongly with values of 0.58 for the West Coast, 0.66 for the East Coast, and 0.69 for the Gulf Coast.

Finally, rarefaction curves further highlighted differences in species accumulation among coasts (Fig. 6). The West Coast curve increased steeply and did not approach an asymptote, indicating continued species accumulation across the sampled range. The East Coast curve rose more gradually and began to level at higher sampling depths, while the Gulf Coast curve increased modestly and reached an apparent plateau at approximately 40 species. Confidence intervals varied across sampling depths, with the narrowest intervals associated with the Gulf Coast curve.

Discussion

The primary goal of this study was to assess whether research-grade iNaturalist observations can be used to characterize large-scale patterns of polychaete diversity across U.S. coastal regions and to identify the strengths and limitations of such data for marine biodiversity research. Polychaete annelids represent one of the most diverse and taxonomically challenging animal groups, with tens of thousands of described species and substantial cryptic diversity, yet they remain underrepresented in large-scale biodiversity syntheses. By aggregating and standardizing research-grade iNaturalist records collected

between 2008 and 2025, this study provides the first continental-scale assessment of polychaete observations derived from a citizen-science platform.

Biogeographic structure and regional turnover

Across all analyses, strong regional differentiation among the East, Gulf and West Coasts emerged as a consistent pattern. Species overlap among coasts was limited, with only a small number of taxa recorded across all three regions. This result aligns with long-standing marine biogeographic frameworks recognizing the Atlantic, Gulf of Mexico, and Pacific coasts as distinct faunal provinces shaped by historical barriers, oceanographic regimes, and habitat structure (Spalding et al., 2007). The limited overlap observed here is therefore biologically plausible and consistent with expectations for benthic invertebrates with regional endemism and constrained dispersal.

At the same time, the pronounced number of species recorded exclusively on the West Coast reflects not only genuine Pacific diversity but also the disproportionate sampling intensity and observer engagement in that region. The Pacific coast supports extensive and readily accessible rocky intertidal habitats, together with long-standing networks of intertidal monitoring programs, community-science initiatives, marine naturalists, and specialist taxonomists that may facilitate both the documentation and identification of coastal biodiversity. In contrast, the East and Gulf coasts are dominated by soft-sediment environments where many polychaetes are burrowing, fouling or otherwise difficult to detect photographically, resulting in lower observed richness despite likely substantial undocumented diversity.

Observation dominance and sampling inequality

Rank-observation curves revealed strongly uneven representation within the iNaturalist dataset. Across all three coasts, a small number of frequently recorded species accounted for a disproportionate share of observations, whereas most species were represented by relatively few records. Such distributions are consistent with opportunistically assembled biodiversity datasets, in which recorded frequencies reflect not only species occurrence but also variation in observer effort, accessibility, detectability, observer preferences, and identification probability (see Di Cecco et al., 2021). The particularly long tail observed for the Pacific coast indicates that its substantially larger observation pool included many species recorded only rarely. This pattern may partly reflect greater sampling intensity, which increases the likelihood of documenting uncommon or infrequently encountered taxa, although differences in regional species richness, habitat diversity, and observer behavior may also contribute.

The contrast between Shannon and Simpson diversity indices further illustrates this dynamic. Higher Shannon diversity on the West Coast reflects the accumulation of rare species through extensive sampling, whereas similar Simpson diversity values across coasts indicate broadly comparable dominance structures driven by a few highly observed taxa. Pielou's evenness revealed the lowest values on the West Coast, consistent with a strong observation skew toward dominant species accompanied by a large pool of rare detections. Together, these metrics demonstrate that differences among regions are

driven primarily by the detection of rare species rather than by fundamentally different dominance patterns.

Limits of Observed Richness

Rarefaction analyses underscore the importance of effort-standardized approaches when interpreting citizen-science biodiversity data. Despite more than 10,000 observations, the West Coast rarefaction curve did not approach an asymptote, indicating that a substantial portion of the regional species pool remains unsampled even under intense observer activity. In contrast, the apparent saturation of the Gulf Coast curve reflects limited sampling rather than true completeness, highlighting the risk of misinterpreting low observed richness as ecological impoverishment. These results reinforce a critical point for marine biodiversity studies using iNaturalist: raw species counts are insufficient for meaningful comparison across regions with vastly different sampling intensity. Instead, rarefaction and complementary analyses are essential for contextualizing observed diversity and avoiding misleading conclusions.

Comparison with other iNaturalist-based biodiversity studies

The patterns observed here are consistent with findings from other studies that have leveraged iNaturalist data to assess biodiversity patterns across taxonomic groups and regions. Previous work using iNaturalist has successfully documented species richness gradients, range expansions, and community composition for birds, plants, insects, and select marine taxa (Koch et al., 2021; Roberts et al., 2022), while simultaneously highlighting biases related to accessibility, observer distribution, and taxonomic expertise (Di Cecco et al., 2021; Cretois et al., 2021). Similar to those studies, the present analysis demonstrates that iNaturalist data can recover dominant and conspicuous components of the observation record and may reveal broad spatial patterns consistent with known biogeographic structure, but underrepresent cryptic, small-bodied, or sediment-dwelling species.

Importantly, this study extends the application of iNaturalist data to a taxonomically complex marine invertebrate group that one would assume would be poorly suited for citizen-science approaches. The ability to recover coherent biogeographic signals and meaningful diversity patterns for polychaetes suggest that the utility of iNaturalist data for marine biodiversity research may be broader than previously assumed, particularly when analyses are framed explicitly around observation patterns rather than true abundance.

A notable feature of the West Coast dataset was the substantial contribution of expert identifiers, including professional polychaete taxonomists. The concentration of expert identifications in this region likely enhanced taxonomic resolution and confidence, enabling the inclusion of species that might otherwise remain unidentified or lumped at higher taxonomic levels. This highlights an often-underappreciated strength of citizen-science platforms: the integration of professional expertise with public participation. Although the east coast does have a history of professional marine biodiversity

monitoring program such as the New England and Mid-Atlantic Rapid Assessment Surveys (Pederson et al., 2021; Fowler et al., 2025), encouraging expert engagement could substantially improve the taxonomic breadth observations. Strategic initiatives, including targeted bioblitzes, partnerships with museums and academic institutions, and outreach to taxonomic specialists, may help reduce regional disparities in data quality and coverage.

Taxonomic interpretation of iNaturalist records

A central implication of this study is that community-science records for polychaetes require explicit taxonomic interpretation. Research-grade iNaturalist observations provide georeferenced, community-vetted records, but they do not substitute for voucher-based determinations, microscopy, molecular data, or comparison with type and museum material. This distinction is particularly important for polychaetes, where species boundaries are often complicated by cryptic diversity, incomplete diagnoses, historical synonymy, and geographically broad species concepts. For example, two of the eight taxa that showed overlap on all three coasts (*F. enigmaticus* and *A. succinea*) are known to contain cryptic lineages that have yet to be sorted taxonomically (Styan et al., 2017; Ferreira et al., 2024). Consequently, the patterns reported here are best interpreted as patterns in documented taxonomic labels and observation effort, rather than as a definitive inventory of U.S. coastal polychaete diversity.

Moving forward: limitations and implications for global change

Despite its strengths, this study also illustrates the inherent limitations of citizen-science datasets for marine biodiversity research. Observation-based records are influenced by habitat accessibility, organism visibility, observer interest, and photographic constraints, all of which bias detection toward larger, more conspicuous taxa. Soft-sediment communities, nocturnal species, and morphologically cryptic taxa remain substantially underrepresented. Moreover, observation frequency should not be interpreted as ecological abundance, and must be framed accordingly. Nevertheless, these limitations do not negate the value of iNaturalist data. Rather, they emphasize the importance of careful analytical framing, explicit acknowledgement of bias, and the use of complementary approaches such as rarefaction and diversity partitioning. When used appropriately, citizen-science data provide an unparalleled opportunity to document biodiversity patterns at spatial and temporal scales that are otherwise unattainable through traditional survey methods alone.

In the context of accelerating marine biodiversity loss and a global taxonomic impediment, platforms like iNaturalist also offer a powerful, if imperfect, tool for biodiversity documentation. While they cannot replace systematic sampling or formal taxonomic revision, citizen-science observations can serve as early warning systems for range shifts, invasions, and local extirpations, and can help prioritize regions and taxa for target research. This study demonstrates that even for a taxonomically challenging group such as polychaetes, citizen-science data can reveal meaningful biogeographic patterns and highlight critical gaps in sampling effort. Moving forward, integrating citizen science with traditional biodiversity

surveys, molecular approaches, and expert taxonomic input will be essential for documenting and conserving marine biodiversity in an era of rapid environmental change.

Declarations

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Funding

There is no funding to declare.

Author Contribution

AD contributed to conception and design, data collection and material preparation. AD and LH both contributed to analysis and writing of the first draft of the manuscript. All authors have approved of this version of the manuscript.

Data Availability

All Python scripts used for mining and analyses along with raw datasets are available via GitHub (https://github.com/parasiteguy/iNaturalist-Polychaete-Biodiversity-Workflow)

References

1. Cahill, A. E., Meglecz, E., & Chenuil, A. (2024). Scientific history, biogeography, and biological traits predict presence of cryptic or overlooked species. *Biological Reviews*, *99*, 546–561. https://doi.org/10.1111/brv.13034
2. Chandler, M., See, L., Copas, K., Bonde, A. M. Z., Lopez, B. C., Danielsen, F., Legind, J. K., Masinde, S., Miller-Rushing, A. J., Newman, G., Rosemartin, A., & Turak, E. (2017). Contribution of citizen science towards international biodiversity monitoring. *Biological Conservation*, *213*, 280–294. https://doi.org/10.1016/j.biocon.2016.09.004
3. Costello, M. J., & Chaudhary, C. (2017). Marine biodiversity, biogeography, deep-sea gradients, and conservation. *Current Biology*, *27*. https://doi.org/10.1016/j.cub.2017.04.060. PR511 - PR527.
4. Cretois, B., Simmonds, E. G., Linnell, J. D. C., van Moorter, B., Rolandsen, C. M., Solberg, E. J., Strand, O., Gundersen, V., Roer, O., & Rod, J. K. (2021). Identifying and correcting spatial bias in opportunistic citizen science data for wild ungulates in Norway. *Ecology and Evolution*, *11*, 15191–15204. https://doi.org/10.1002/ece3.8200

5. Cull, B. (2025). Online crowdsourced data from iNaturalist can assist monitoring of invasive mosquitoes. *Insects*, *16*, 128. https://doi.org/10.3390/insects16020128
6. Davinack, A. A., & Fitzhugh, K. (2025). Cryptic diversity does not justify weakening systematics standards. A comment on: 'Adapting practices to accelerate the scientific description of invertebrate cryptic species', (2025) by Vivien *et al. Biology Letters, 22*, 20250652. https://doi.org/10.1098/rsbl.2025.0652
7. Davinack, A. A., Carlton, J. T., & Pederson, J. (2024). Annotated checklist and community composition of introduced, cryptogenic, and native polychaetes in floating dock communities of New England, USA. *Journal of the Marine Biological Association of the United Kingdom*, *104*, e103. https://doi.org/10.1017/S0025315424000961
8. Di Cecco, G. J., Barve, V., Belitz, M. W., Stucky, B. J., Guralnick, R. P., & Hurlbert, A. H. (2021). Observing the observers: how participants contribute data to iNaturalist and implications for biodiversity science. *BioScience*, *71*, 1179–1188. https://doi.org/10.1093/biosci/biab093
9. Dickinson, J. L., Shirk, J., Bonter, D., Bonney, R., Crain, R. L., Martin, J., Phillips, T., & Purcell, K. (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, *10*, 291–297. https://doi.org/10.1890/110236
10. Dinnage, R., Grady, E., Neal, N., Deck, J., Denny, E., Walls, R., Seltzer, C., Guralnick, R., & Li, D. (2025). PhenoVision: A framework for automating and delivering research-ready plant phenology data from field images. *Methods in Ecology and Evolution*, *16*, 1763–1780. https://doi.org/10.1111/2041-210X.70081
11. Engel, M. S., Ceriaco, L. M. P., Daniel, G. M., Dellape, P. M., Lobl, I., Marinov, M., Reis, R. E., Young, M. T., Dubois, A., Agarwal, I., & Zacharie, C. K. (2021). The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. *Zoological Journal of the Linnean Society*, *193*, 381–387. https://doi.org/10.1093/zoolinnean/zlab072
12. Ferreira, L. M., Sexias, V. C., Villalobos-Guerrero, T. F., & Santos, C. S. G. (2024). Another candidate to human-mediated transport: molecular evidence confirms that *Alitta succinea* (Nereididae: Annelida) is a widespread species. *Marine Biodiversity*, *54*, 12. https://doi.org/10.1007/s12526-024-01407-x
13. Fowler, A. E., Blakeslee, A. M. H., Davinack, A., Aguilar, R., Andersen, M., Benadon, C., Choong, H. H. C., Green-Gavrielidis, L., Greenberg, S. R., Hartshorn, E., Hobbs, N-V., Labbe, S., Larson, K., McCuller, M., Moloney, D. M., Parsons, S. K., Stancil, C., Thornber, C., Pederson, J., & Carlton, J. T. (2025). Caribbean Creep meets Chesapeake Creep: marine bioinvasions and community shifts along the Mid-Atlantic Coast, USA. *Biological Invasions*, *27*, 192. https://doi.org/10.1007/s10530-025-03646-w
14. Hending, D. (2024). Cryptic species conservation: a review. *Biological Reviews*, *100*, 258–274. https://doi.org/10.1111/brv.13139
15. Herrera, D. J., Schalk, C. M., Jensen, A. J., Goldstein, B. R., Rooney, B. R., Kays, R., McShea, W. J., & Cove, M. V. (2025). iNaturalist and structured mammal surveys reflect similar species richness but capture different species pools across the United States. *Ecology and Evolution*, *25*, e71805. https://doi.org/10.1002/ece3.71805

16. IPBES (2019). *Global assessment report on biodiversity and ecosystem services*. E.S. Brondizio, J. Settele, S. Diaz, & H.T. Ngo (Eds.). Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany. https://doi.org/10.5281/zenodo.3831673
17. Jaureguiberry, P., Titeux, N., Wiemers, M., Bowler, D. E., Coscieme, L., Golden, A. S., Guerra, C. A., Jacob, U., Takahashi, Y., Settele, J., Diaz, S., Molnar, Z., & Purvis, A. (2022). The direct drivers of recent global anthropogenic biodiversity loss. *Science Advances*, *8*, eabm9982. https://doi.org/10.1126/sciadv.abm9982
18. Koch, J. B. U., Tabor, J. A., Montoya-Aiona, K., & Eiben, J. A. (2021). The invasion of *Megachile policaris* (Hymenoptera: Megachilidae) to Hawai'i. *Journal of Insect Science*, *21*, 1–9. https://doi.org/10.1093/jisesa/ieab065
19. Marin-Gomez, O. H., Flores, C. R., & del Arizmendi, C., M (2022). Assessing ecological interactions in urban areas using citizen science: Insights from hummingbird-plant meta-networks in a tropical megacity. *Urban Forestry & Urban Greening*, *74*, 127658. https://doi.org/10.1016/j.ufug.2022.127658
20. Marsh, D. M., Finnegan, Holmes, C., & Li, J. (2026). Comparing iNaturalist and traditional surveys for estimating richness, diversity and abundance of reptiles and amphibians. *Herpetologica*, *82*, 1–11. https://doi.org/10.1655/Herpetologica-D-24-00056
21. Mason, B. M., Mesaglio, T., Heitmann, J. B., Chandler, M., Chowdhury, S., Gorta, S. B. Z., Grattarola, F., Groom, Q., Hitchcock, C., Hoskins, L., Lowe, S. K., Marquis, M., Pernat, N., Shirey, V., Baasanmunkh, S., & Callaghan, C. T. (2025). iNaturalist accelerates biodiversity research. *BioScience*, *75*, 953–965. https://doi.org/10.1093/biosci/biaf104
22. Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., & Worm, B. (2011). How many species are there on earth and in the ocean? *PLoS Biology*, *9*, e1001127. https://doi.org/10.1371/journal.pbio.1001127
23. Nygren, A. (2014). Cryptic polychaete diversity: a review. *Zoologica Scripta*, *43*, 172–183. https://doi.org/10.1111/zsc.12044
24. Paknia, O., Sh, H. R., & Koch, A. (2015). Lack of well-maintained natural history collections and taxonomists in megadiverse developing countries hampers global biodiversity exploration. *Organisms Diversity & Evolution*, *15*, 619–620. https://doi.org/10.1007/s13127-015-0202-1
25. Pederson, J., Carlton, J. T., Bastidas, C., David, A., Grady, S., Green-Gavrielidis, L., Hobbs, N-V., Kennedy, C., Knack, J., McCuller, M., O'Brien, B., Osborne, K., Pankey, S., & Trott, T. (2021). 2019 Rapid Assessment Survey of marine bioinvasions of southern New England and New York, USA, with an overview of new records and range expansions. *BioInvasions Records*, *10*, 227–237. https://doi.org/10.3391/bir.2021.10.2.01
26. Potgieter, L. J., ter Huurne, M. B., & Richardson, D. M. (2024). Community science can inform invasive species management: *Melaleuca* (Myrtaceae) in South Africa. *Ecological Solutions and Evidence*, *5*, e12391. https://doi.org/10.1002/2688-8319.12391
27. Ranasinghe, U. G. S. L., Eberle, J., Thormann, J., Bohacz, C., Benjamin, S. P., & Ahrens, D. (2022). Multiple species delimitation approaches with COI barcodes poorly fit each other and

morphospecies - an integrative taxonomy case of Sri Lankan Sericini chafers (Coleoptera: Scarabaeidae). *Ecology and Evolution*, *12*, e8942. http://dx.doi.org/10.1002/ece3.8942

28. Roberts, C. J., Verges, A., Callaghan, C. T., & Poore, A. G. B. (2022). Many cameras make light work: opportunistic photographs of rare species in iNaturalist complement structured surveys of reef fish to better understand species richness. *Biodiversity and Conservation*, *31*, 1407–1425. https://doi.org/10.1007/s10531-022-02398-6
29. Rosa, R. M., Cavallari, D. C., & Salvador, R. B. (2022). iNaturalist as a tool in the study of tropical molluscs. *Plos One*, *17*, e0268048. https://doi.org/10.1371/journal.pone.0268048
30. Rouse, G., Pleijel, F., & Tilic, E. (2022). *Annelida*. Oxford University Press.
31. Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdana, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A., Lourie, S. A., Martin, K. D., McManus, E., Molnar, J., Recchia, C. A., & Robertson, J. (2007). Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *BioScience*, *57*, 573–583. https://doi.org/10.1641/B570707
32. Struck, T. H., & Cerca, J. (2019). Cryptic species and their evolutionary significance. *Encyclopedia of Life Sciences*, 1–9. https://doi.org/10.1002/9780470015902.a0028292
33. Styan, C. A., McCluskey, C. F., Sun, Y., & Kupriyanova, E. K. (2017). Cryptic sympatric species across the Australian range of the global estuarine invader *Ficopomatus enigmaticus* (Fauvel, 1923) (Serpulidae, Annelida). *Aquatic Invasions*, *12*, 53–65. https://doi.org/10.3391/ai.2017.12.1.06
34. Szabolcs, M., Wenner, B., Mizsei, E., & Lengyel, S. (2026). Distribution and sampling biases of citizen science observations uploaded to iNaturalist from Hungary. *Citizen Science: Theory and Practice*, *11*, 15–15. https://doi.org/10.5334/cstp.861
35. Theobald, E. J., Ettinger, A. K., Burgess, H. K., DeBey, L. B., Schmidt, N. R., Froehlich, H. E., Wagner, C., HilleRisLambers, J., Tewksbury, J., Harsch, M. A., & Parrish, J. K. (2015). Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. *Biological Conservation*, *181*, 236–244. https://doi.org/10.1016/j.biocon.2014.10.021
36. Vivien, R., Martin, P., Pawlowski, J., & Alther, R. (2025). Adapting practices to accelerate the scientific description of invertebrate cryptic species. *Biology Letters*, *21*, 20250385. http://dx.doi.org/10.1098/rsbl.2025.0385
37. Wagner, M. A., Wang, Z., Anthony, C. J., Rivera, E. A., & Strader, M. E. (2025). iNaturalist data suggests a recent northward expansion of the upside-down jellyfish, *Cassiopea*. *Science of the Total Environment, 992*, 179948. https://doi.org/10.1016/j.scitotenv.2025.179948
38. Ward, D., Melbourne-Thomas, J., Pecl, G. T., Evans, K., Green, M., McCormack, P. C., Novaglio, C., Trebilco, R., Bax, N., Braiser, M. J., Cavan, E. L., Edgar, G., Hunt, H. L., Jansen, J., Jones, R., Lea, M-A., Makomere, R., Mull, C., Semmens, J. M., Shaw, J., Tinch, D., van Steveninck, T. J., & Layton, C. (2022). Safeguarding marine life: conservation of biodiversity and ecosystems. *Reviews in Fish Biology and Fisheries*, *32*, 65–100. https://doi.org/10.1007/s11160-022-09700-3

Figures



Figure 1

Geographic distribution of research-grade polychaete (Annelida: Polychaeta) observations from iNaturalist used in this study.

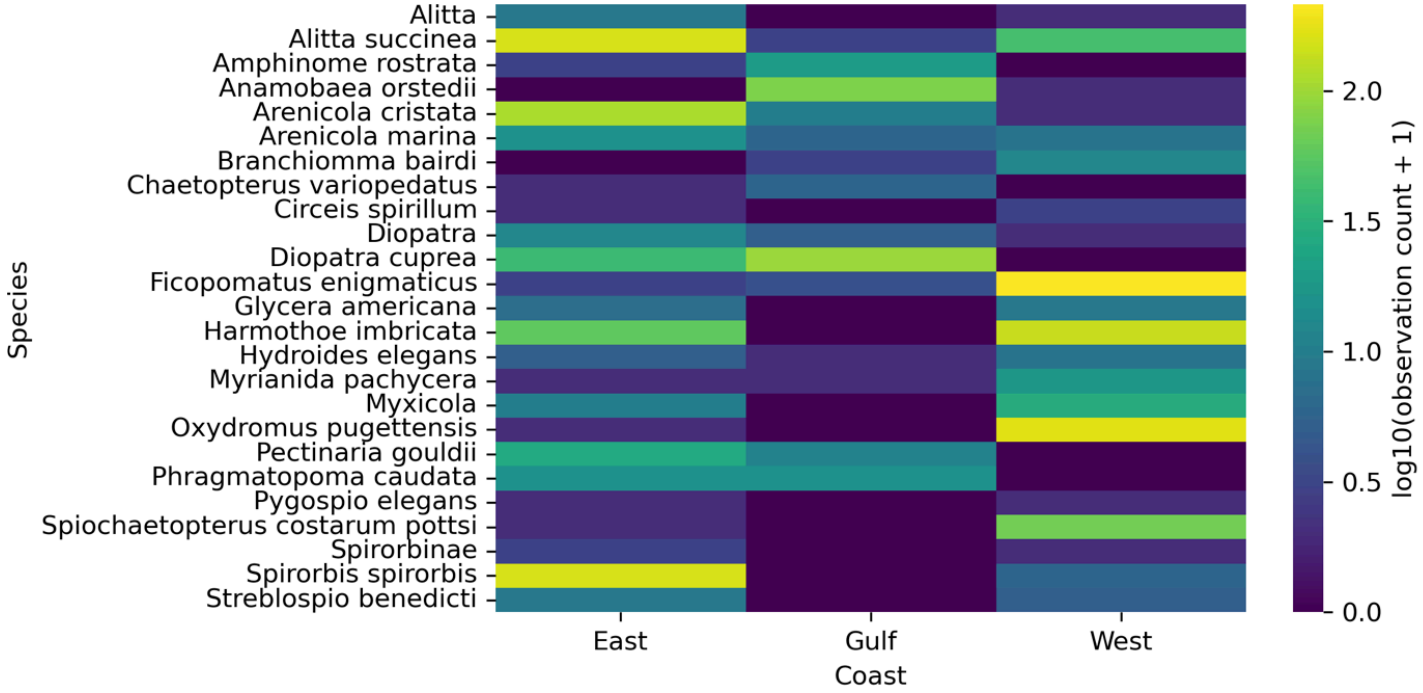


Figure 2

Heatmap of pairwise species overlap among polychaete assemblages recorded along the East, Gulf, and West Coasts of the United States using research-grade iNaturalist observations. Cell values indicate the number of species shared between each pair of coastal regions, with darker shading corresponding to greater overlap.

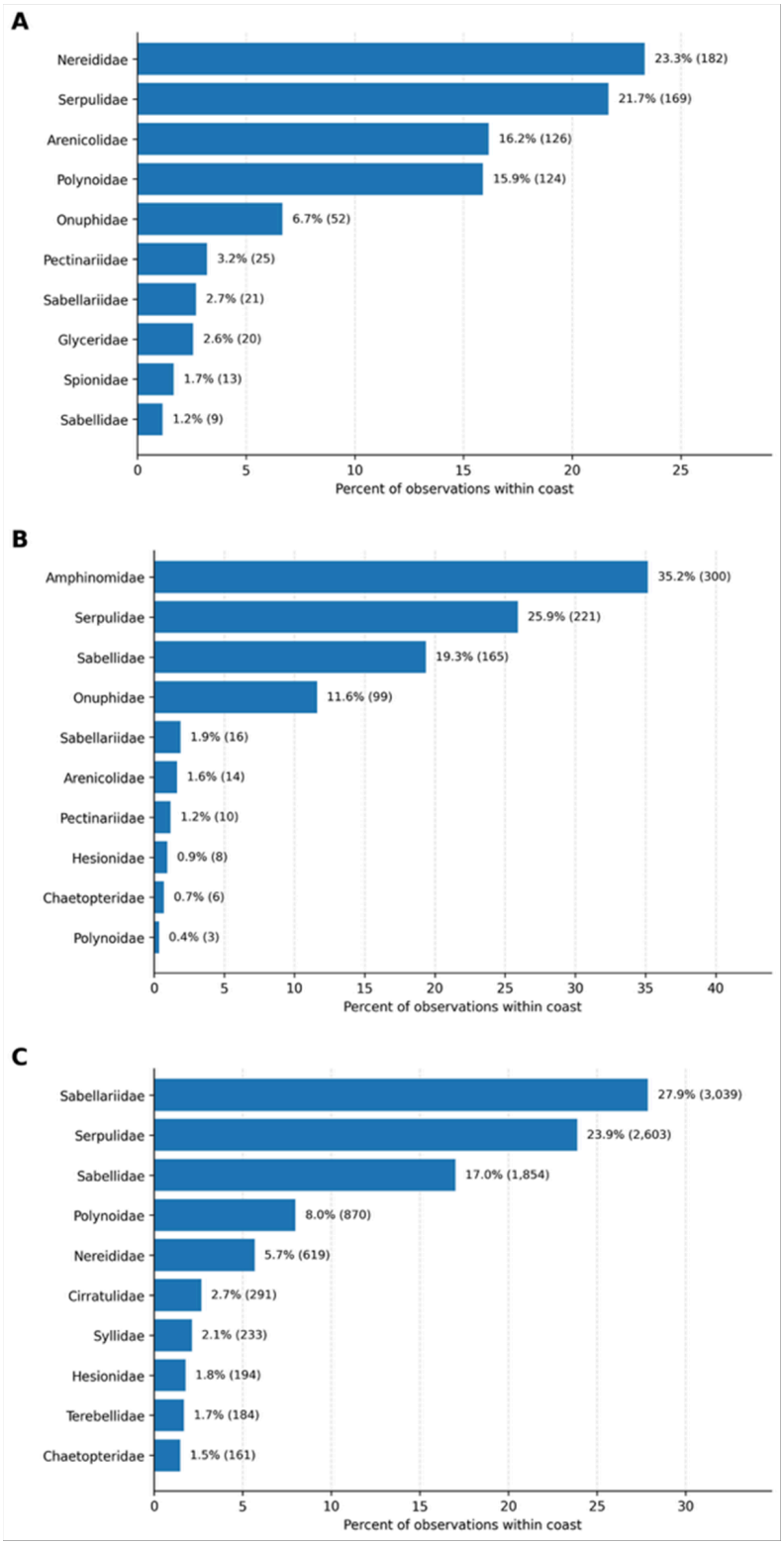


Figure 3

Top 10 most frequent polychaete families recorded on A) West Coast, B) Gulf Coast and C) East Coast.

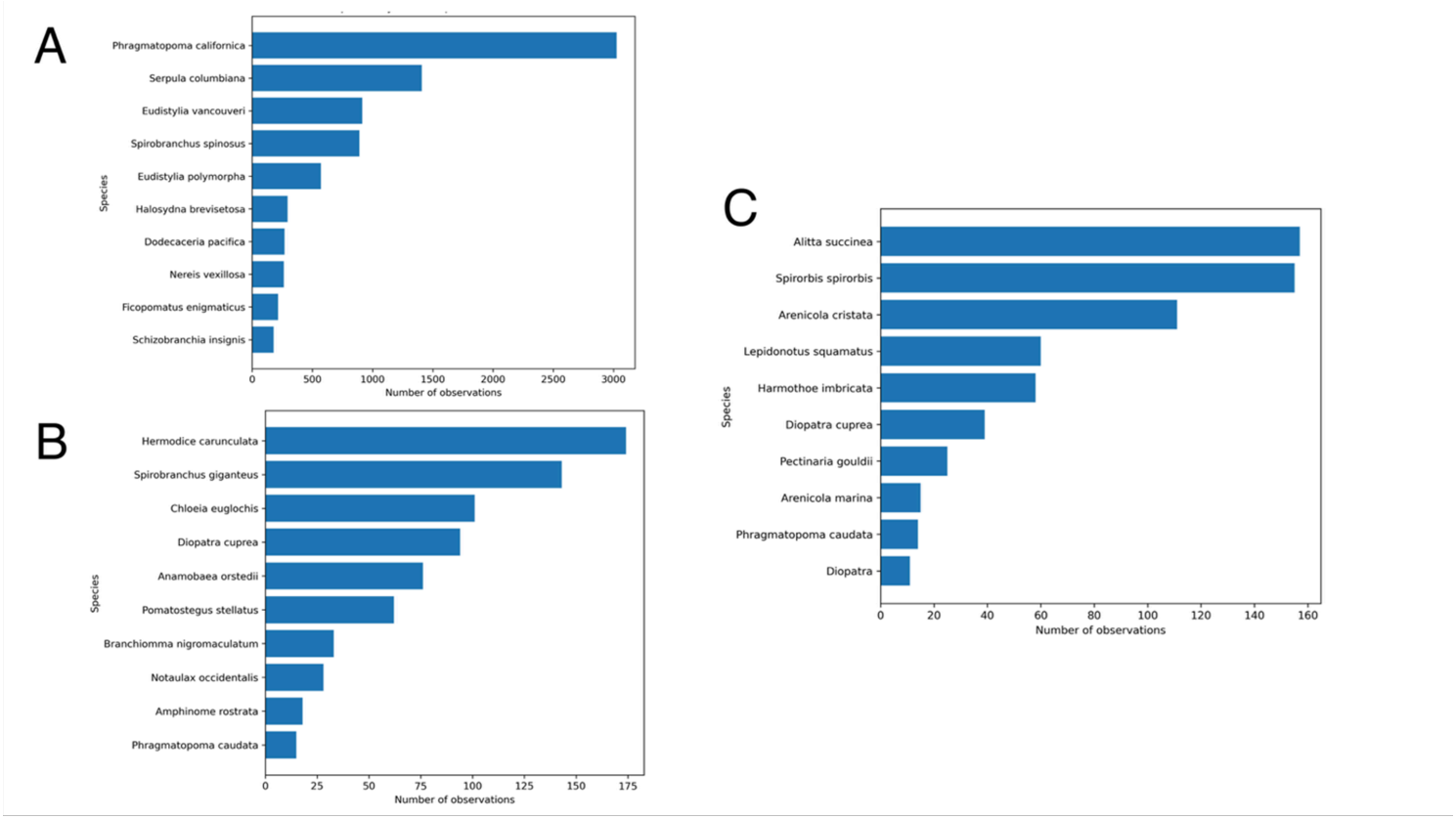


Figure 4

Top 10 most frequent polychaete species recorded on A) West Coast, B) Gulf Coast, and C) East Coast.

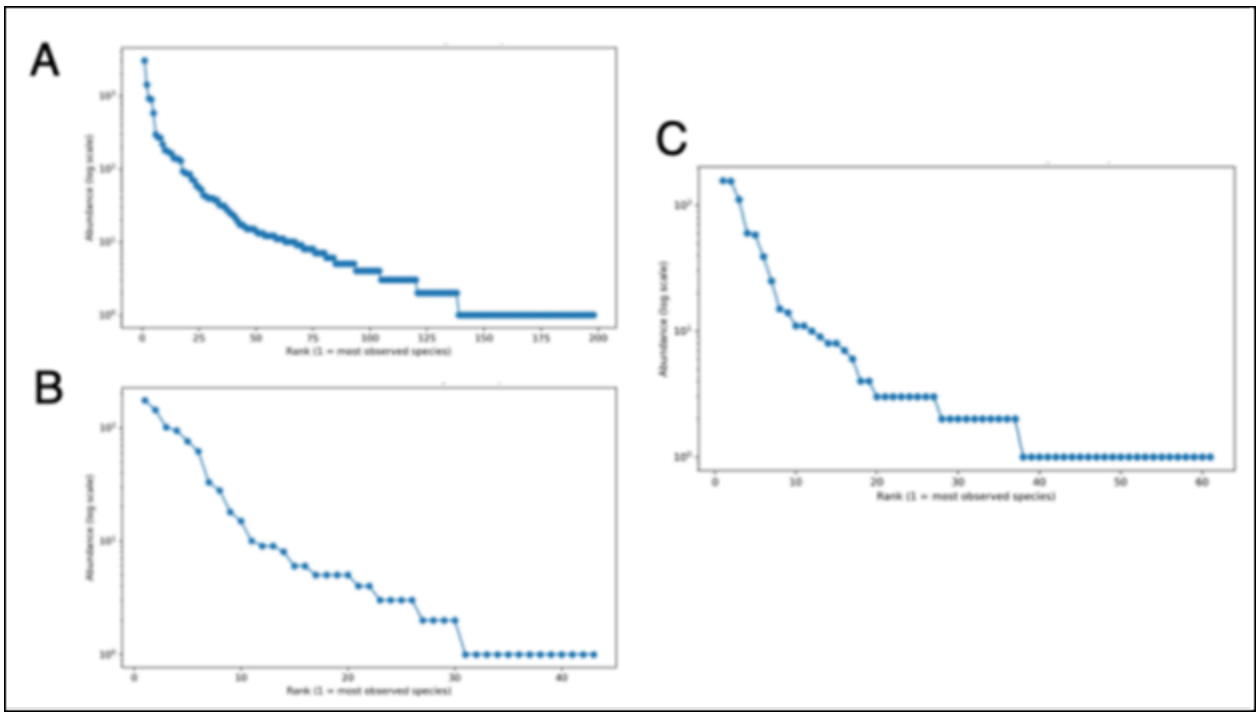


Figure 5

Rank-observation curves for polychaete species observed along the West, Gulf and East Coasts of the United States based on research-grade iNaturalist records collected between 2008 and 2025. Species are ranked in descending order of observation frequency within each coastal region, with each point representing a single species. Observation counts are plotted on a logarithmic scale to illustrate dominance patterns and the distribution of rare taxa. Differences in curve length reflect variation in observed species richness among coasts, while differences in slope indicate variation in evenness and sampling intensity across regimes.

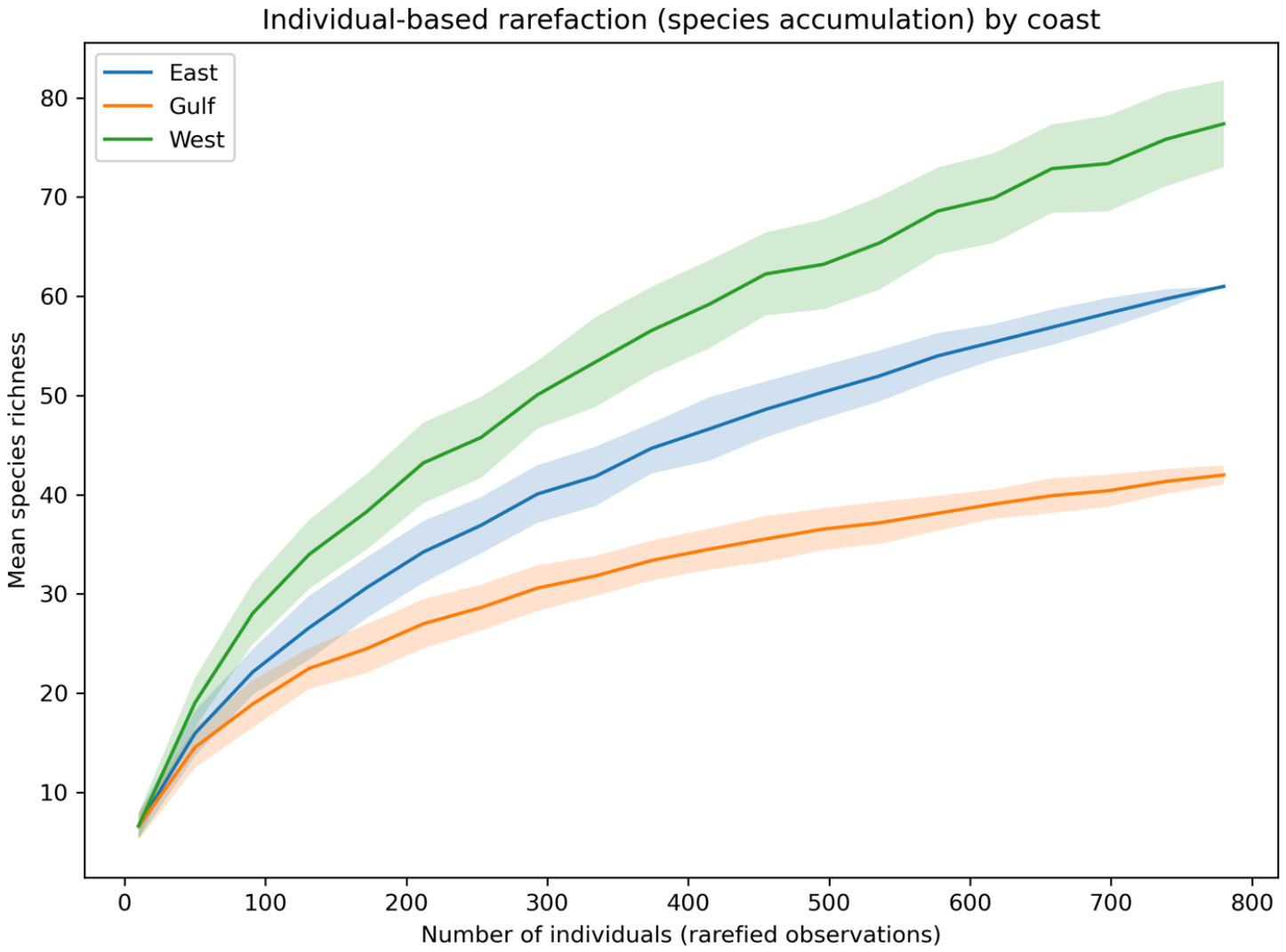


Figure 6

Individual-based rarefaction (species accumulation) curves for polychaete species observed along the East, Gulf, and West Coasts of the United States using research-grade iNaturalist records collected between 2008 and 2025. Curves show the mean expected species richness as a function of the number of randomly subsampled observations for each coastal region. Shaded areas represent ±1 standard deviation across randomization replicates. Differences among curves reflect variation in observed richness and sampling intensity rather than standardized estimates of true regional diversity.