Check for updates

Molecular Ecology Resources

WILEY

MOLECULAR ECOLOGY RESOURCES

RESOURCE ARTICLE OPEN ACCESS

Secrets From the Shelf: Museomics Unlocks Molecular Data From Historical Type Specimens of Octocorals

Catherine S. McFadden[1] | Declan Morrissey[2] | Claudia F. Vaga[2] | Samuel A. Vohsen[2] | Kaveh Samimi-Namin[3] | Teresa Cerqueira[2,4] | Kaela A. Coil[2] | Elyssia S. Gonzalez[2,5] | Laura Macrina[6] | Makiri Sei[2] | Herman H. Wirshing[2] | Yehuda Benayahu[7] | Andrea M. Quattrini[2]

[1]Department of Biology, Harvey Mudd College, Claremont, California, USA | [2]Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA | [3]Marine Evolution and Ecology Group, Naturalis Biodiversity Center, Leiden, the Netherlands | [4]Institute of Marine Sciences – Okeanos, University of the Azores, Horta, Portugal | [5]Ocean, Coastal, and Earth Sciences Program, School of Earth, Environmental, and Marine Sciences, University of Texas Rio Grande Valley, Brownsville, Texas, USA | [6]Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia | [7]School of Zoology, George S. Wise Faculty of Life Science, Tel Aviv University, Tel Aviv, Israel

Correspondence: Catherine S. McFadden (mcfadden@hmc.edu)

Received: 7 July 2026 | **Revised:** 6 October 2026 | **Accepted:** 30 October 2026

Keywords: DNA barcoding | genome skimming | mitogenome | Octocorallia | Symbiodiniaceae | target capture | UCEs | ultraconserved element

ABSTRACT

Name-bearing type specimens are the primary, authoritative references that anchor the application of species names and provide the morphological basis for assessing species identity. As molecular data are increasingly used to identify and discriminate species, it has become necessary to genetically characterize type specimens to ensure that correct names are applied to genetic lineages. Many existing type specimens are, however, hundreds of years old and preserved under conditions that may degrade or destroy DNA. To test the feasibility of obtaining genomic data from historical type material of octocorals, we extracted DNA from dry- and fluid-preserved tissues of 571 historical specimens, including type material and non-type vouchers, collected from 3 to 200 years ago and sequenced them using either genome-skimming (i.e., low-coverage shotgun sequencing) or target-capture of ultraconserved elements and exons (UCEs). We recovered complete, circularized mitochondrial genomes from > 40% of specimens plus mitochondrial barcode sequences (*mtMutS*, *COI*) from > 60%, hundreds of UCE loci, and *28S rDNA* barcodes from > 75%. Preservation was a more important predictor of recovery of genomic data than age of the collection; among specimens > 50 years old, genome-skimming of dry-preserved material yielded the most and highest quality sequence data. Incorporating sequence data from types into reference databases and molecular phylogenies will improve the accuracy of DNA-based species identification and help resolve the identity of confusing and morphologically challenging taxa, including species considered to be *incertae sedis* (of unknown genus or family) or *taxa inquirenda* (names of uncertain application).

1 | Introduction

Species, defined as evolutionarily independent lineages under a unified species concept (de Queiroz 2007), are the fundamental units of evolutionary biology. Delineation of species (i.e., recognition and discrimination of independent lineages) is necessary to study processes such as reproductive biology, genetic connectivity, physiology and adaptation and increasingly to effectively manage populations for conservation (Coates et al. 2018). The naming of species and consistent application of those names over space and time are important, allowing information learned by different researchers working at different locations and different

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Molecular Ecology Resources* published by John Wiley & Sons Ltd.

times to be attributed to the same evolutionarily independent lineage. Misidentifications and misapplications of names have important consequences for our understanding of geographic ranges, habitat specificity, population sizes and the basic biology of species (e.g., Costa et al. 2015; Rodrigues et al. 2022; Vázquez-Ramírez et al. 2025).

The modern practice of taxonomy (i.e., the naming of species) established by Linnaeus (1758) has changed relatively little over the past two centuries. As currently governed by the International Code of Zoological Nomenclature, species are formally recognized based on designated, name-bearing type specimens, exemplars that remain the primary, authoritative references for the identity of that species (ICZN 1999). Correct application of a species' name to additional specimens has historically required that they conform to the morphological description of the type material. As molecular systematic studies conducted over the past several decades have revealed the existence of cryptic species complexes—cases in which evolutionarily independent lineages cannot be distinguished by morphology alone—it has become increasingly necessary to use genetic data to discriminate species and to apply correct names (Bickford et al. 2007; Zhang and Wiens 2026). Consequently, it has also become necessary to genetically characterize type material, a challenge when existing type specimens may be hundreds of years old and preserved under conditions that compromise or destroy DNA (Hahn et al. 2021).

Although the importance of genetically characterizing type specimens is widely recognized (Letch et al. 2025), technical limitations have hindered progress. Early attempts to obtain genetic data from historical type specimens yielded limited success because traditional PCR-based barcoding approaches perform poorly with fragmented and/or chemically modified DNA typical of older museum specimens. Recent advances in high-throughput sequencing now provide new opportunities to recover genomic information from challenging museum material (Raxworthy and Smith 2021). Genomic data from name-bearing type specimens can integrate these historically important specimens into modern molecular systematic analyses, enabling the confident application of species names and establishing genomic reference data for future studies of biodiversity and evolution. Genomic data have been obtained successfully from name-bearing types up to 200 years old in groups such as vertebrates (Scherz et al. 2020; Straube et al. 2021; Castañeda-Rico et al. 2022) and insects (Ahrens et al. 2026), while success rates with small, soft-bodied invertebrates have been more variable (Palandačić et al. 2024).

Octocorallia is a class of anthozoan cnidarians that encompasses approximately 3600 nominal species of corals including soft corals, sea fans and whips (gorgonians), sea pens, and other forms characterized by polyps that feature eight, usually pinnate, tentacles and eight internal mesenteries (McFadden et al. 2026). Unlike the more familiar reef-building stony corals, only a handful of octocoral species secrete a massive carbonate skeleton that retains its integrity upon death of the organism. Instead, most species have sclerites, free skeletal elements of crystalline calcite that are embedded in their tissues to impart rigidity to an otherwise hydrostatic skeleton. The shapes, sizes and anatomical distribution of these sclerites are the primary taxonomic characters that have historically been used to distinguish species of octocorals (Fabricius and Alderslade 2001). The complex and highly variable shapes of sclerites are, however, difficult to quantify objectively (Vargas et al. 2010; Carlo et al. 2011); their size distributions may be influenced by environmental factors (West 1997); and some taxa lack sclerites entirely (Benayahu et al. 2017; McFadden et al. 2022). Increasingly, integrative taxonomic efforts have revealed independent genetic lineages (i.e., species) that are not morphologically distinct (McFadden et al. 2017) or can only be distinguished by morphological characters that had not been considered important historically (or simply could not be imaged effectively in the past) and are therefore not presented in the original species descriptions (Alderslade 2001; McFadden et al. 2009; Benayahu et al. 2018; Samimi-Namin et al. 2026). Determining if these newly recognized lineages have been described previously and, if so, applying correct names to them, requires redescription of the original type material to confirm morphology and, if possible, obtain diagnostic molecular characters for the species. In cases where molecular data cannot be obtained from historical types, the collection of topotypic specimens (i.e., new material from the type locality that matches the original morphological description of a species) from which genetic data can be obtained is an increasingly common practice (e.g., Bridge et al. 2024). Although topotypes provide an important alternative when historical types cannot be sequenced, they cannot fully replace genomic data obtained directly from name-bearing type specimens (Letch et al. 2025).

Recent advances in DNA extraction and library preparation protocols have allowed the successful recovery of genomic data from type specimens of octocorals ranging in age from 20 to 174 years (Morrissey et al. 2026). These promising results have raised several important questions whose answers will help guide efforts to obtain genomic data from additional types and to establish best practices for preserving newly collected type material: (i) the relationships between specimen age (i.e., date of collection), method of preservation, and successful recovery of genomic data; (ii) whether genome-skimming (i.e., low-coverage shotgun sequencing) or target capture is the more effective approach for obtaining sequence data from degraded historical material (Call 2020); and (iii) the relationships between specimen age, preservation, sequencing protocol, and the recovery of contaminant sequences, either from the endosymbionts that occur naturally in corals or from external sources associated with long-term storage and handling of museum specimens. To address these questions, we extracted DNA from the preserved tissue of 571 octocoral type specimens and other historically important material collected from 3 to 200 years ago, and used either a genome-skimming or target capture approach to obtain genomic data. Success was evaluated as numbers and median lengths of ultraconserved elements and exons (collectively called UCEs herein; Quattrini et al. 2018, 2024) recovered, as well as complete or partial mitochondrial genomes, loci used for barcoding of octocorals (*mtMutS*, *COI* and *28S rDNA*), and numbers of off-target (i.e., non-octocoral) sequence reads. Off-target sequences were identified as either naturally occurring endosymbionts (i.e., Symbiodiniaceae) or exogenous contaminants (e.g., human, fungal, bacterial). Here, we present the sequence data obtained, make publicly available valuable genomic resources for name-bearing types, and explore the

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

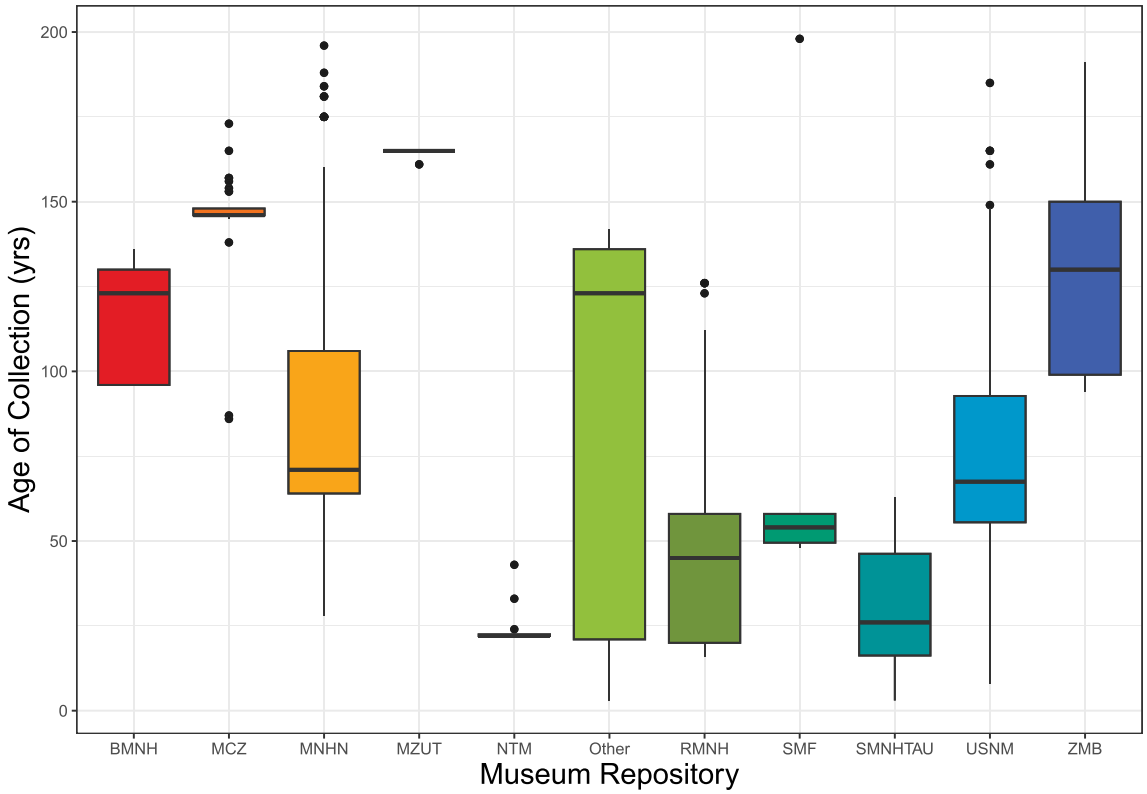


FIGURE 1 | Age of octocoral type specimens sequenced from different museum repositories. BMNH: British Museum of Natural History, London; MCZ: Museum of Comparative Zoology, Harvard University; MNHN: Muséum National d'Histoire Naturelle, Paris; MZUT: Museo Regionale di Scienze Naturali, Turin, Italy; NTM: Museum and Art Gallery of the Northern Territory, Australia; RMNH: Naturalis Biodiversity Center, Netherlands (formerly Rijksmuseum van Natuurlijke Historie, including the collections of the former Zoological Museum of Amsterdam, ZMA); SMF: Senckenberg Nature Museum, Frankfurt; SMNHTAU: Steinhardt Museum of Natural History, Tel Aviv University; USNM: Smithsonian National Museum of Natural History, US; ZMB: Museum für Naturkunde, Berlin. 'Other' includes all repositories from which fewer than 5 specimens were sequenced: Musée Océanographique de Monaco (MOM); Museo de Zoología, Universidad de Costa Rica (MZUCR); Museum of Natural History, Vienna (NHMW); Queensland Museum, Australia (QM); Seto Marine Biological Laboratory, Japan (SMBL); Tasmanian Museum and Art Gallery, Australia (TMAG); Florida Museum of Natural History, US (UF); and Zoologisches Museum Hamburg (ZMH).

effects of collection age, method of preservation (e.g., ethanol, formalin, or dried) and sequencing approach on the recovery of genomic data from legacy collections of octocorals. The approaches and outcomes we report are widely applicable to historical collections of other soft-bodied marine organisms and preserved soft tissues.

2 | Methods

2.1 | Collections

Octocorals were sampled from the collections of 18 natural history museums that house original type material (Figure 1). Types were selected based on their relevance to other ongoing projects, prioritizing shallow-water zooxanthellate soft corals (families Cladiellidae, Sarcophytidae, Xeniidae) and deep-water North Atlantic taxa (e.g., Acanthogorgiidae, Alcyoniidae, Primnoidae). Both name-bearing types (holotypes, lectotypes, syntypes) and type series (paratypes, paralectotypes) were sampled as well as some historically important non-type specimens, such as vouchers examined by past octocoral experts (e.g., Deichmann 1936; Table S1). Specimens were collected 3–200 years ago (median collection age = 66 years), and the majority were preserved dry or in 70% ethanol (Figure 2). A few specimens were known to have been fixed in formalin prior to long-term storage in ethanol, but the original fixative could not be verified for most material. To validate the identity of each type and its status, the details on the labels were verified with the collection database and in some cases even with old handwritten museum catalogues. Each colony was morphologically compared to the original description of the type material, including when possible the colony shape and the shapes and sizes of sclerites. In some cases, however, only small fragments of types remained or were available for examination, precluding verification of all morphological characters. In cases where exact collection dates were missing from museum records or the published literature, we estimated the collection date of specimens based on historical information (e.g., known dates of expeditions) or used the date of publication of the species description as the latest possible collection date.

2.2 | DNA Extraction and Sequencing

DNA was extracted using a Qiagen DNeasy Blood & Tissue kit, with a few modifications to the manufacturer's published

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

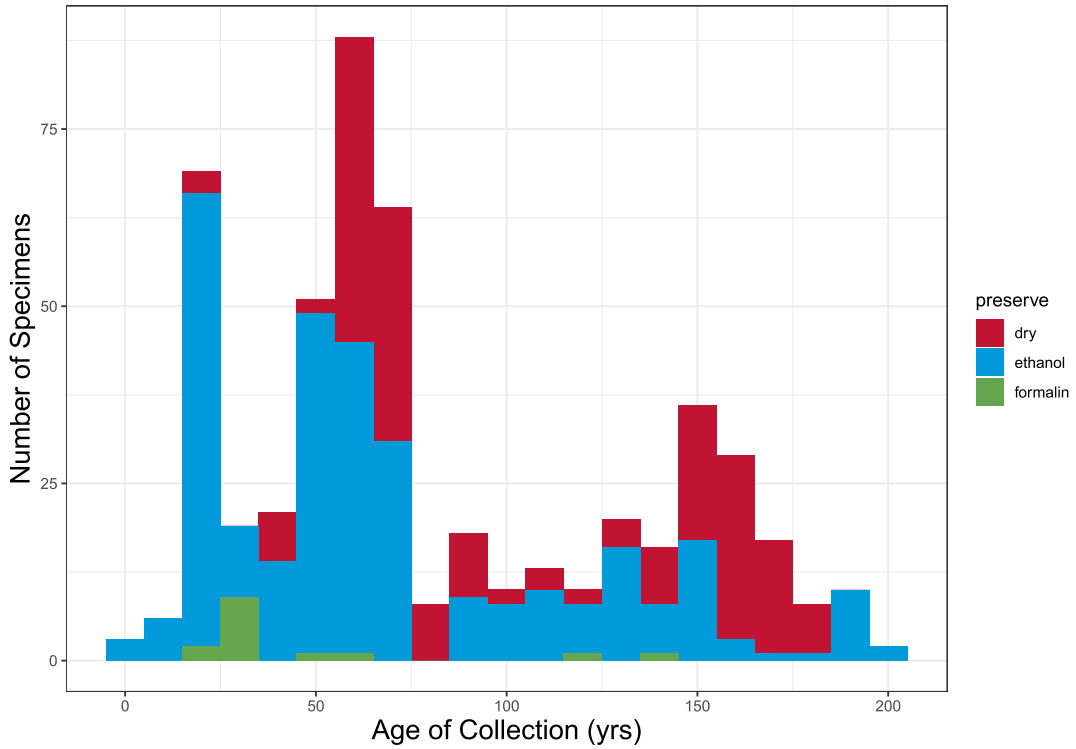


FIGURE 2 | Collection age and method of preservation of octocoral type material sequenced for this study.

protocol (Morrissey et al. 2026). Approximately 25 mg of ethanol-preserved or dry coral tissue was minced with a sterile scalpel blade on a clean surface (aluminium foil or paper towel) and added to 180 μL buffer ATL. After addition of proteinase K (20 μL), tissue was incubated at 56°C for 48 h (or two overnights), with a second aliquot of 20 μL proteinase K added after 24 h. After lysis and washing steps, DNA was eluted from the spin column twice, each time with 50 μL buffer AE or molecular grade water. For most samples collected > 20 years ago, DNA was then cleaned using a Zymo Clean & Concentrator-10 kit and eluted from the wash column with 35 μL heated (70°C) Elution Buffer (EB). Clean DNA was quantified with a Qubit fluorometer using the Invitrogen High Sensitivity quantification assay.

DNA was either sent to a commercial facility (Arbor Biosystems, Ann Arbor, MI; University of Oregon Genomics and Cell Characterization Core Facility, UO G3CF) for library preparation or prepared at the Smithsonian Laboratories of Analytical Biology (LAB). Either a NEBNext Ultra II DNA Library Kit or Watchmaker Genomics DNA Library Prep Kit was used to prepare samples for genome-skimming. The former was used with some modifications as in Morrissey et al. (2026): briefly, the reaction volume was reduced by half, iTru Y-yoke oligonucleotide adaptors (Glenn et al. 2019) were used instead of NEBNext Adaptors, KAPA Pure Beads ratio was increased to 1.2×, adaptor ligation time was 30 min, iTru i5 and i7 indices (Glenn et al. 2019) were used, and 10–15 cycles of PCR enrichment were conducted. For target capture sequencing, libraries were prepared as described in Quattrini et al. (2018) using a Kapa Hyper Prep Kit (Kapa Biosystems) with dual-indexed iTru adapters. MYBAITS protocol v.4 (Arbor Biosystems) was used to target and enrich pools of 8 libraries using the octocoral-v2 bait set of Erickson et al. (2021). Libraries were sequenced on an Illumina NovaSeq 6000, targeting 20–30 M reads for genome skimming and 4–5 M PE reads for target capture (150 bp PE reads).

2.3 | Sequence Analysis

Sequences were trimmed to remove adaptors and low-quality reads using fastp v.1.1.0 (Chen et al. 2018). Mitochondrial (mt) genomes were assembled and recovered using MitoFinder v1.4.2 (Allio et al. 2020), which performs de novo assembly of trimmed reads and identifies mitochondrial contigs through similarity searches against a curated database of reference mitogenomes (in this case, previously published octocoral mt genomes in GenBank). All complete, circularized mitogenomes recovered were annotated using Geneious v. 2026.0.1 (https://www.geneious.com/) with a curated database of reference mitogenomes and fine-tuned using the Open Reading Frame tool using the transl_table 4 within Geneious. Annotated mitogenomes were submitted to GenBank (Table S1). When a complete, circularizable mitogenome was not recovered, any mt genes that were identified by Mitofinder were confirmed to be octocoral using NCBI BLAST (Camacho et al. 2009). Mitochondrial regions most commonly used for DNA barcoding of octocorals (~750 bp 5′ end of *mtMutS*; ~800 bp 5′ end of *COI*) were aligned to reference files (e.g., McFadden et al. 2011; Quattrini et al. 2024) to check for annotation errors and then submitted to GenBank (Table S1). Sequences for all other mitochondrial protein-coding genes (PCGs) identified by Mitofinder are available on dryad (see Data Availability).

MITObim v. 1.9.1 (Hahn et al. 2013) was used to assemble *28S rDNA* from the trimmed sequences, using a seed file that consisted of a published *28S rDNA* sequence (complete or partial)

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from the same genus. The identity of *28S* sequences was confirmed using NCBI BLAST, and the ~650 bp region commonly used as a DNA barcode for octocorals (McFadden and van Ofwegen 2013; McCartin et al. 2024) was aligned to reference files (Quattrini et al. 2024; McCartin et al. 2026), annotated and submitted to GenBank (Table S1).

Trimmed sequences were assembled into contigs using SPAdes v. 3.15.5 (Bankevich et al. 2012). Phyluce v 1.7.3 (Faircloth 2016) was used to identify ultraconserved and exon loci (hereafter, UCEs) by matching probes (octocoral v2 bait set, Erickson et al. 2021) to contigs (*phyluce_assembly_match_contigs_to_probes*) with a minimum coverage of 70% and minimum identity of 70%. Loci were extracted as fastas (*phyluce_assembly_get_fastas_from_match_counts*) and the number and median length of loci recovered from each individual were summarized (*phyluce_assembly_get_fasta_lengths*). Fastas were then aligned with a reference dataset (Quattrini et al. 2024) using MAFFT v. 7.130b (Katoh and Standley 2013) (*phyluce_ align_seqcap_ align*), internally trimmed using gblocks (Castresana 2000) (*phyluce_align_get_ gblocks_trimmed_ alignments_from_untrimmed*, -b1 0.5, -b2 −0.5, -b3 10, -b4 5) and concatenated into a data matrix with 50% or 75% of taxa present for each locus (*phyluce_align_get_only_loci_with_min_taxa*).

To confirm the correct phylogenetic placement of sequenced type specimens, alignments of UCEs and single barcode loci (*mtMutS*, *28S rDNA*) with taxon-comprehensive reference files (e.g., Quattrini et al. 2024) were used to construct maximum likelihood phylogenies in IQtree v. 2.2.1 (Minh et al. 2020) with ultrafast bootstrapping (-bb 1000, Hoang et al. 2018). A partitioned model (-p) was used for the UCE dataset. The best model of nucleotide substitution for each barcode gene or UCE partition was found with ModelFinder (-m TESTMERGE, Kalyaanamoorthy et al. 2017). For UCEs, separate phylogenetic analyses were conducted for: (1) all samples from which > 5 UCE loci were recovered; (2) only samples with > 100 UCE loci; (3) samples with < 100 but > 5 UCEs that were subset to reconstruct selected families or particular clades (e.g., only Sarcophytidae, only Acanthogorgiidae, only Scleralcyonacea).

For 23 samples that were initially recovered in a clade sister to the Octocorallia, we then used NCBI's BLASTn to detect octocoral UCE loci. We BLASTed the unaligned loci generated by *phyluce_assembly_explode_get_fastas_file* against a combined database containing 32 octocoral genomes downloaded from NCBI's Genome Resource (code available on dryad, see Data Availability). These loci were then extracted and put back into the alignment at the *phyluce_align_seqcap_align* step.

2.4 | Identification of Symbiont and Contaminant Sequences

Mitochondrial or *28S rDNA* sequences derived from endosymbiont taxa or contaminants were detected using NCBI BLAST and classified as fungal, dinoflagellate (i.e., Symbiodiniaceae), human or other (typically another eukaryote). In addition, trimmed reads were classified with Kraken2 classify v. 2.1.3 (Wood et al. 2019) using a custom database comprising the standard reference genomes (human, bacteria, archaea, viruses, plasmids) as well as eight fungal genomes, 32 octocoral genomes, and 12 Symbiodiniaceae genomes retrieved from GenBank (Table S2). Database construction followed manual instructions with prepending NCBI taxids to the fasta headers of added genomes. The number of off-target reads was calculated as the sum of reads classified as Symbiodiniaceae, human, fungi, archaea, bacteria, or virus. Finally, phyloFlash v3.4.2 (Gruber-Vodicka et al. 2020) was used to assemble SSU rRNA genes of any off-target sources of DNA.

2.5 | Statistical Analyses

We used inferential analyses to address three questions: (1) whether sequencing success varied with collection age and preservation method; (2) whether museum repository influenced sequencing success and contamination after accounting for collection age, preservation, and sequencing protocol; and (3) whether genome-skimming and target-capture differed in their recovery of ultraconserved element (UCE) loci. Analyses were conducted in R 4.6.0 (R Core Team 2026) on an x86_64-pc-linux-gnu platform in a Google Colab Linux environment; package versions, scripts, metadata, and archived outputs are available in dryad.

Data were analysed using generalized linear models (GLMs) and, where appropriate, generalized linear mixed-effects models (GLMMs), with model family selected according to response type. Because collection age, preservation, and sequencing protocol were not independent, protocol was retained as a core predictor in pooled analyses. Collection age was analysed as a continuous predictor after log transformation. Some specimens were sequenced more than once; repeated observations were therefore considered explicitly in sensitivity analyses.

To test how sequencing success changed with collection age, preservation, and protocol, we fit models evaluating either the interaction between age and protocol or the interaction between age and preservation, while retaining the third variable as an adjustment term. Numbers of UCE loci recovered were analysed with a negative binomial GLM. The number of mitochondrial protein-coding genes (PCGs) was analysed with a beta-binomial GLM because this response is bounded (maximum = 14 genes) and showed greater among-specimen variation than expected under a standard binomial model. Recovery of a circularized mitogenome and *mtMutS* barcode was analysed using logistic regression. Median UCE length was analysed as a continuous response after log transformation. The proportions of reads retained after trimming, octocoral reads, and off-target reads were analysed as proportion-response GLMs. Because fungal, human, and Symbiodiniaceae contamination were sparse and zero-heavy, they were analysed primarily as presence/absence variables rather than raw percentages. Detection of SSU rRNA sequences from specific taxa of contaminants was also analysed as presence/absence variables. Because positive detections of these focal taxa were restricted to genome-skimming samples, preservation-related patterns were evaluated primarily within the genome-skimming subset. An apparent association between detection of human and *Cutibacterium* SSU rRNA was tested using Fisher's exact test. *COI* and *28S* barcode outcomes were analysed in two steps: first, recovery of any barcode sequence

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

versus none; second, among recovered sequences, contaminant versus genuine identity.

Museum effects were tested using adjusted GLMs that included museum repository, log-transformed collection age, preservation, and sequencing protocol. A full age × preservation × museum interaction was not used because the dataset was too sparse and unbalanced across those combinations for reliable inference. Contaminant identity among recovered *28S* sequences was therefore treated mainly as a descriptive outcome rather than a primary inferential response. Protocol effects on UCE recovery were tested in the full dataset using a model including protocol, log-transformed collection age, and preservation, and in the subset of samples sequenced by both methods using a paired Wilcoxon signed-rank test. Because a subset of specimens was represented by more than one sequencing run, we also performed GLMM sensitivity analyses for key binary outcomes by including specimen identity as a random intercept. Model fit was assessed using standard residual diagnostics, including simulation-based residual checks for the main non-Gaussian models.

To further test whether the relative performance of genome-skimming vs. target-capture protocols depended on specimen age, we conducted an age-class sensitivity analysis. Specimens were assigned to two age classes, <50 years and ≥ 50 years, to maintain adequate sample sizes in both sequencing-protocol groups. For each response variable, we fit models including sequencing protocol, age class, and their interaction, with preservation method and museum repository retained as adjustment terms where estimable. Number of UCE loci recovered was analysed using a negative binomial GLMM, median UCE length was analysed after log-transformation, number of mitochondrial PCGs was analysed as a beta-binomial response, and circularized mitogenome, *mtMutS*, *COI* and *28S* recovery were analysed as binomial responses. Post hoc contrasts were used to compare genome-skimming and target-capture within each age class. As a secondary sensitivity analysis where sample size allowed, the same age-class comparison was repeated for ethanol-preserved specimens only.

3 | Results

3.1 | Sequencing Success

Sequence data were obtained from 519 of the 571 specimens from which DNA was extracted (collection age 3–198 years, median = 66 years; Table S1, Figure 2). Library preparation failed for an additional 40 specimens (collection age 27–200 years, median = 106 years), and too little DNA was recovered from 12 specimens to proceed with library preparation (collection age 32–175 years, median = 75 years). Libraries were successfully prepared and sequenced using <1 ng to 250 ng of total DNA (median = 10.1 ng), resulting in 0 to >656 M PE reads per sample (median = 17 M PE reads). Larger quantities of DNA generally yielded higher numbers of trimmed reads, UCEs and mitochondrial protein-coding genes (Figure S1).

Historical specimens from 36 families of octocorals were successfully sequenced, the majority from order Malacalcyonacea ($n = 459$) and family Sarcophytidae ($n = 213$). Other families that were well-represented were Cladiellidae ($n = 76$), Acanthogorgiidae ($n = 54$), Xeniidae ($n = 29$), Gorgoniidae ($n = 18$), Plexauridae ($n = 15$), Alcyoniidae ($n = 11$), and from the order Scleralcyonacea the families Primnoidae ($n = 19$) and Ellisellidae ($n = 8$). The largest number of sequenced specimens came from the collections of MNHN ($n = 164$) followed by RMNH/ZMA ($n = 75$), USNM ($n = 74$) and SMNHTAU ($n = 62$). The average collection age of type specimens differed among museums, with most of the oldest material (median age of collection >100 years) coming from BMNH, MCZ, MZUT and ZMB, and younger material (median age of collection <50 years) from NTM, RMNH and SMNHTAU (Figure 1). The majority of sequenced specimens were ethanol-preserved ($n = 311$) followed by dry ($n = 193$). Only 15 specimens could be confirmed to have been fixed in formalin. Most dry-preserved specimens were collected more than 50 years ago, while almost all recently collected material (<50 years) had been preserved in ethanol (Figure 2). Fleshy soft corals (e.g., Sarcophytidae, Cladiellidae) and gorgonians (i.e., octocorals with a supporting axis) were both well-represented among the dry material.

3.2 | Recovery of Loci

Some specimens ($n = 35$) were sequenced more than once using different library preparation protocols and sequencing approaches, for a total of 554 sets of sequence data obtained (Table S1). Numbers of PE150 sequence reads retained after trimming to remove adaptors and low-quality sequences ranged from 505 to 1.35×10^7 (median = 2,588,410.5) for target-captured samples and 0–1.22×10^8 (median = 23,887,332.5) for genome-skimmed samples (Figure 3a). UCE loci were recovered from 543 of 554 samples. The median number of UCE loci recovered across all samples was 635 (range = 0–2490), although fewer than 100 loci were recovered from 181 samples (Figure 3b). Complete, circularized mitogenomes were assembled from 230 samples, and complete or partial *mtMutS* and octocoral *COI* barcode sequences from 338 and 339 samples respectively (Figure 4). Complete or partial *28S rDNA* barcodes were recovered from 416 samples. Mitogenomes and barcodes were recovered disproportionately from material that had been genome-skimmed compared to target-captured (Figures 3d, and 4). Complete mitogenomes, *mtMutS* and *COI* barcodes were obtained from 49% to 70% of genome-skimmed types, but from only 16% (mitogenome), 30% (*mtMutS*) and 40% (*COI*) of target-captured types (Figure 4a–c), while *28S rDNA* barcodes were recovered from 78% of genome-skimmed and 65% of target-captured samples (Figure 4d).

3.3 | Effects of Protocol, Age and Preservation on Sequencing Success

Recovery of genomic data depended strongly on sequencing protocol and preservation, while the effect of specimen age differed among preservation classes for several core outcomes. These three factors were non-independent because target-capture was used disproportionately to sequence more recent collections (<50 years), and recent collections were disproportionately ethanol-preserved (Figure 2). In models testing age-by-protocol

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

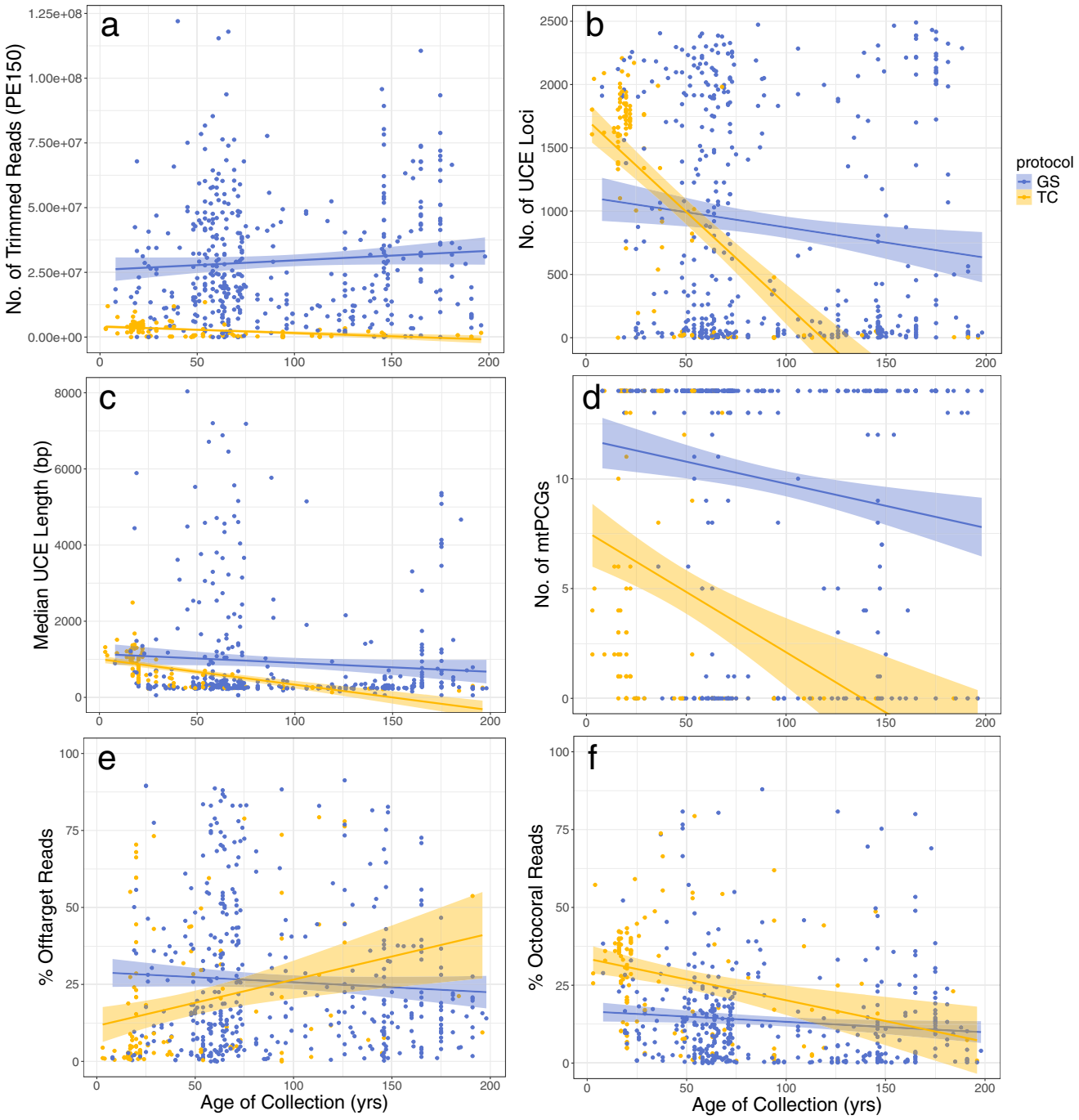


FIGURE 3 | Effect of specimen age and sequencing protocol on number of sequence reads and loci recovered from octocoral type specimens. GS = genome-skimming (blue), TC = target-capture (gold). (a) number of reads (PE150) remaining after trimming; (b) number of UCE loci recovered; (c) median length (bp) of UCE loci; (d) number of mitochondrial protein-coding genes (mtPCGs) recovered; (e) percent of off-target reads; (f) percent of octocoral reads.

effects while adjusting for preservation, both numbers of UCEs recovered and median UCE length showed significant age × protocol interactions (negative binomial GLMM for UCE recovery, likelihood-ratio statistic = 38.838, $p = 4.61 \times 10^{-10}$; Gaussian GLMM for log-transformed median UCE length, likelihood-ratio statistic = 4.573, $p = 0.0325$; Table S3), with genome-skimming recovering more and longer UCE loci from older collections than target capture (Figure 3b,c). Recovery of mitochondrial PCGs varied significantly with preservation (beta-binomial GLMM, likelihood-ratio statistic = 28.488, $p = 6.51 \times 10^{-7}$; Table S3) but did not show a significant age × protocol interaction. In addition, significantly more UCE loci were recovered by genome-skimming than target capture from 15 specimens (median collection age = 94 years) that were prepared in the same facility (LAB) using both methods (paired Wilcoxon signed-rank test, $V = 103$, $p = 0.0017$; Table S4A,B; Figure S2). Target capture recovered very few UCE loci from these samples (median = 6, maximum = 51), whereas genome-skimming recovered an order of magnitude more loci (median = 54, maximum = 701).

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

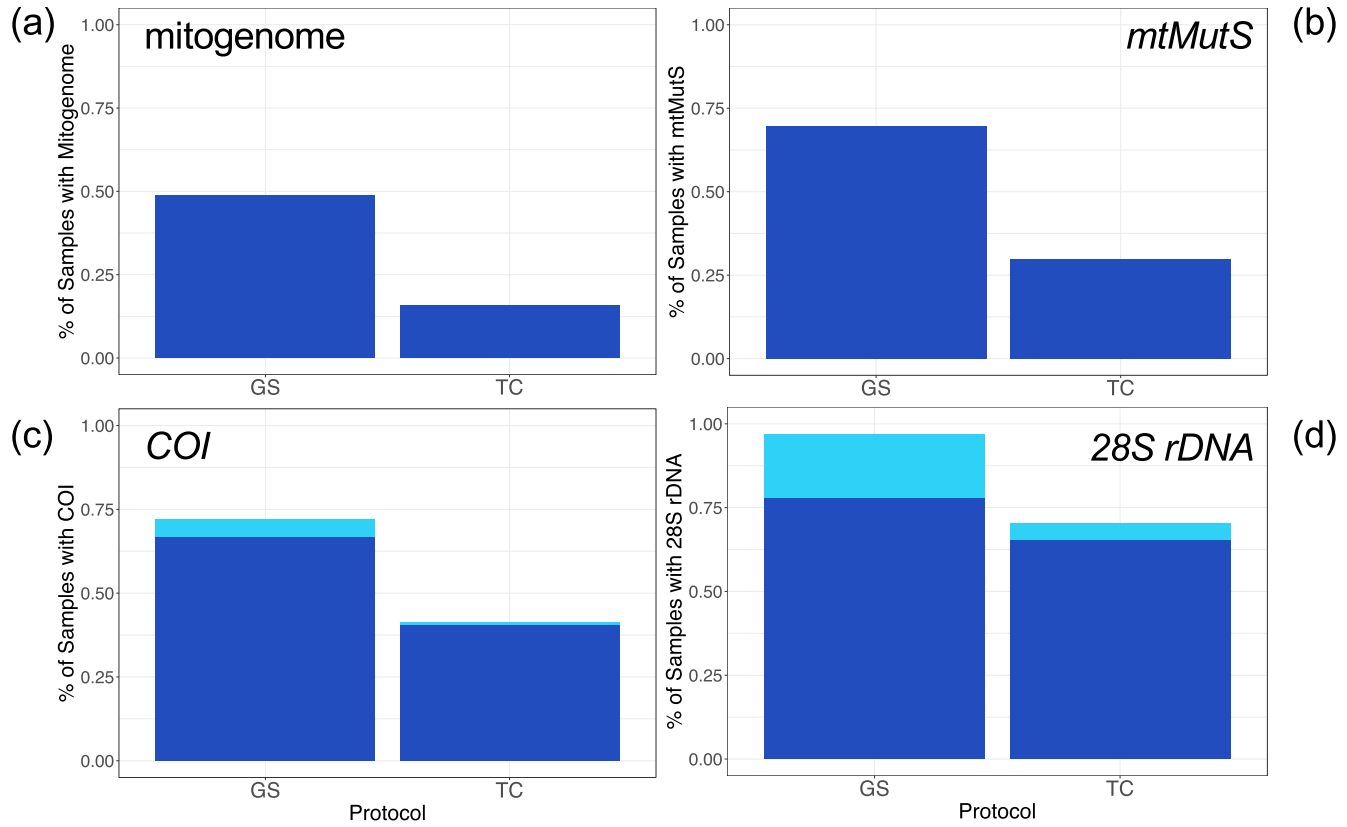


FIGURE 4 | Percent of octocoral type specimens from which (a) complete mitogenomes, (b) *mtMutS* barcode, (c) *COI* barcode and (d) *28S rDNA* barcode sequences were recovered from genome-skimmed (GS) ($n = 433$) vs. target-captured (TC) ($n = 121$) samples. Lighter blue = contaminant sequences.

The age-class sensitivity analysis supported a more nuanced interpretation of protocol performance. In the full dataset, the protocol × age-class interaction was significant for the number of UCE loci recovered (negative binomial GLMM, likelihood-ratio statistic = 15.331, $p = 9.0 \times 10^{-5}$; Table S5A–C). Genome-skimming and target capture did not differ significantly among specimens < 50 years old (post hoc GS/TC ratio = 1.21, $p = 0.519$), whereas genome-skimming recovered significantly more UCE loci than target capture among specimens ≥ 50 years old (GS/TC ratio = 5.38, $p = 1.60 \times 10^{-10}$). Median UCE length showed a similar age-dependent protocol effect (Gaussian GLMM, protocol × age-class likelihood-ratio statistic = 9.984, $p = 0.00158$; Table S5B): genome-skimming and target capture did not differ significantly among specimens < 50 years old ($p = 0.125$), but genome-skimming recovered longer UCE loci among specimens ≥ 50 years old ($p = 0.00379$). The sensitivity analysis of ethanol-preserved material produced the same qualitative result for recovery of UCE loci, with no significant difference between protocols among specimens < 50 years old but substantially greater recovery of UCE loci by genome-skimming among specimens ≥ 50 years old (Table S5D).

Collection age interacted significantly with preservation to influence several measures of sequence recovery. Significant age × preservation effects were detected for number of UCEs (negative binomial model, age × preservation term LRT = 31.027, $p = 1.83 \times 10^{-7}$), median UCE length (Gaussian model, age × preservation term LRT = 12.577, $p = 0.00186$), and number of PCGs recovered (beta-binomial model, age × preservation term LRT = 17.003, $p = 2.03 \times 10^{-4}$; Table S6). Number of UCEs recovered declined with collection age in ethanol-preserved and formalin-fixed material, whereas dry-preserved material showed little evidence of age-related decline (Figure 5a). Recovery of mitochondrial PCGs likewise declined with age in ethanol-preserved material, while the direction of the formalin effect was similar but should be interpreted cautiously because few specimens could be confidently identified as formalin-fixed (Figure 5c; Table S6). Median UCE length declined with collection age in ethanol-preserved material but showed little evidence of decline in dry-preserved material, which retained relatively longer loci at older ages (Figure 5b; Table S6). Median UCE length was consistently low in formalin-fixed material, although this pattern should also be interpreted cautiously because of the small formalin subset. Similar preservation-dependent age effects were also detected for recovery of circularized mitogenomes (binomial GLMM, age × preservation interaction likelihood-ratio statistic = 14.858, $p = 5.94 \times 10^{-4}$), *mtMutS* barcodes (binomial GLMM, age × preservation interaction likelihood-ratio statistic = 9.359, $p = 0.00928$), and any *COI* barcode sequence (binomial GLMM, age × preservation interaction likelihood-ratio statistic = 7.785, $p = 0.0204$; Table S6).

After adjustment for collection age, preservation, and sequencing protocol, museum repository also remained a significant predictor of sequence recovery. Adjusted museum effects were detected for recovery of UCEs (negative binomial GLMM, likelihood-ratio statistic = 41.976, $p = 7.57 \times 10^{-6}$) and PCGs (beta-binomial GLMM, likelihood-ratio statistic = 52.201,

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

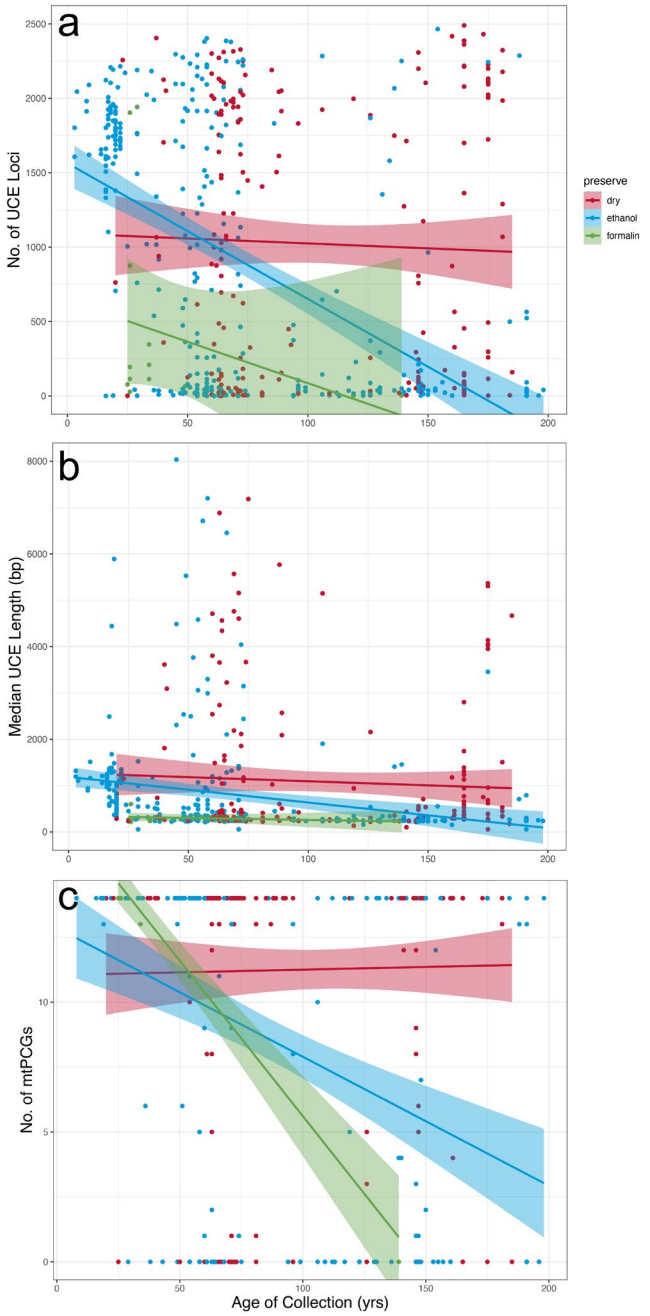


FIGURE 5 | Effect of age of collection and preservation on (a) number of UCE loci recovered; (b) median length (bp) of UCE loci; (c) number of mitochondrial protein-coding genes (mtPCGs) recovered. blue: Ethanol-preserved; red: Dry-preserved; green: Formalin-fixed.

$p = 1.05 \times 10^{-7}$), as well as the fraction of off-target reads (beta GLMM, likelihood-ratio statistic = 31.315, $p = 5.20 \times 10^{-4}$; Table S7). In contrast, adjusted museum effects on fractions of specific contaminant types recovered were weak or not estimable and are therefore treated descriptively below.

3.4 | Contaminant and Off-Target Sequences

Overall, 39.4% (median) of all trimmed sequence reads were classified by Kraken. Among the reads that were classified, genome-skimmed samples contained a higher proportion of off-target (i.e., non-octocoral) than octocoral sequences (median 17.1% off-target vs. 9.3% octocoral), whereas target-captured samples showed the opposite pattern (median 6.9% off-target vs. 28.0% octocoral), reflecting the use of capture baits specifically to enrich octocoral loci. Off-target reads increased and octocoral reads decreased with collection age more strongly in target-captured than in genome-skimmed samples (Figure 3e,f), but the corresponding adjusted age × protocol interactions were weak in the final models (beta GLMM for off-target reads, likelihood-ratio statistic = 3.382, $p = 0.0659$; beta GLMM for octocoral reads, likelihood-ratio statistic = 0.671, $p = 0.4128$; Table S8).

The most common off-target reads belonged to Symbiodiniaceae (median 7.2% of reads in zooxanthellate species; 0.15% in azooxanthellate taxa; Figure 6). The most frequently recovered symbiont genus was *Cladocopium* (median = 51% of all Symbiodiniaceae reads), with some specimens also hosting high frequencies (> 90% of reads) of *Breviolum*, *Durusdinium* and *Symbiodinium* (Table S9). In azooxanthellate species, bacterial sequences were the most common off-target reads (median 3.3%; 0.9% in zooxanthellate taxa). Across all samples, a median of 0.3% of reads were human-derived, while sequences derived from fungi, archaea, and viruses each made up < 0.02% of all reads. Percent of off-target sequences was, however, extremely variable, with some samples having > 75% of reads comprised of Symbiodiniaceae, human, or bacterial sequences (Figure 6). Samples with the highest levels of contamination by human or bacterial DNA (> 25% of reads) had almost all been collected > 50 years ago. In the adjusted models, however, human contamination showed no strong effect of age, protocol, or preservation, whereas fungal contamination showed a protocol effect, with higher probability of detection in target-captured than genome-skimmed samples, although this result should be interpreted cautiously because the model showed convergence warnings (binomial GLMM, likelihood-ratio statistic = 7.984, $p = 0.00472$). The age effect for fungal contamination was weak (likelihood-ratio statistic = 2.959, $p = 0.0854$; Table S8).

Off-target SSU rRNA sequences detected using phyloFlash included taxa ranging from Archaea to other metazoans such as nematodes, molluscs, and crustaceans (Table S10). Taxa that were detected in the highest frequency were human, Symbiodiniaceae, the fungi *Aspergillus* (=*Phialosimplex*) and *Malassezia*, and the human-associated bacteria *Cutibacterium*, *Staphylococcus*, and *Streptococcus* (Tables S1 and S11). Because positive detections of these seven focal taxa were restricted to genome-skimmed samples, preservation-related patterns were evaluated primarily within the genome-skimming subset. Contamination by *Aspergillus* showed strong age and preservation effects; it was detected significantly more often in older material (binomial GLMM, likelihood-ratio statistic = 28.94, $p = 7.46 \times 10^{-8}$) and in dry compared to ethanol-preserved specimens (Wald $z = -4.999$, $p = 5.77 \times 10^{-7}$). *Aspergillus* was detected in 65 of 187 dry-preserved genome-skimmed samples (34.8%) but only 2 of 233 ethanol-preserved genome-skimmed samples (0.9%) and in no formalin-fixed genome-skimmed samples. In contrast, human SSU rRNA was detected more frequently in ethanol-preserved genome-skimmed samples (47 of 233; 20.2%) than in dry-preserved genome-skimmed samples (18 of 187; 9.6%), although neither age nor preservation was statistically significant in the adjusted model (age: likelihood-ratio statistic = 0.631, $p = 0.4268$; preservation: likelihood-ratio

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

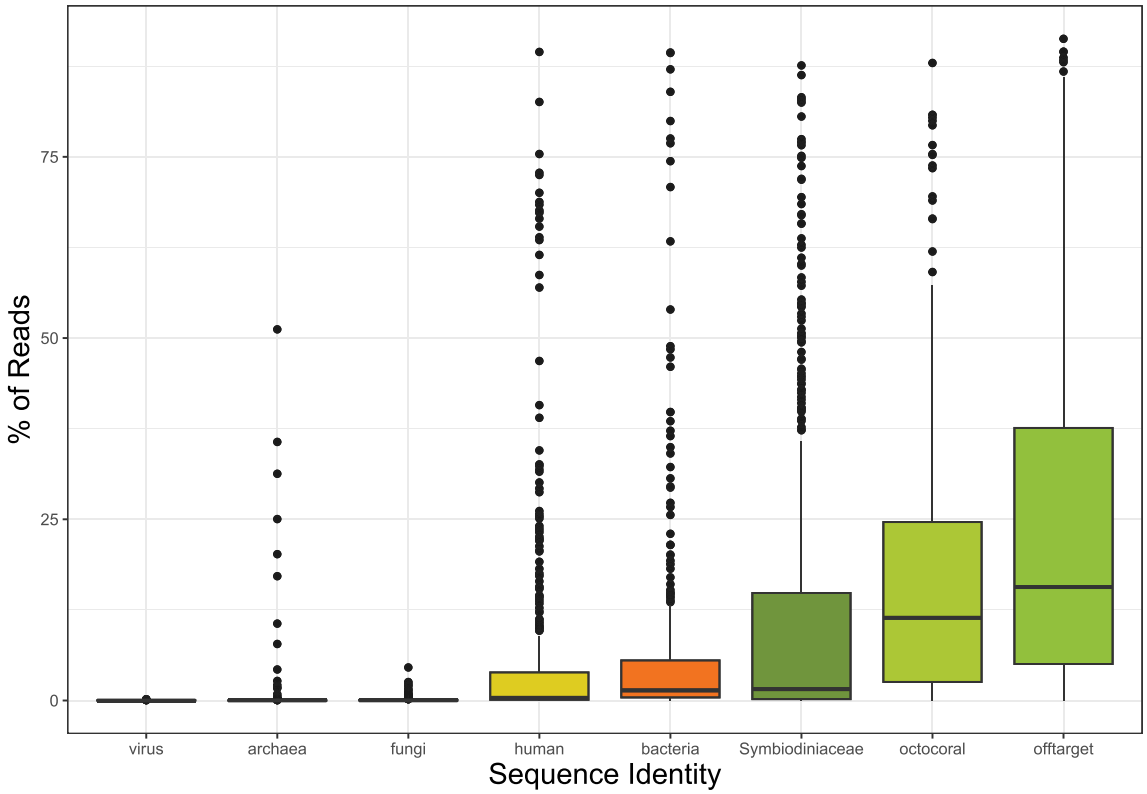


FIGURE 6 | Identity of sequence reads recovered from type specimens of octocorals, classified using Kraken2. Off-target = sum of all non-octocoral reads.

statistic = 0.788, $p = 0.6743$). *Cutibacterium* showed a similar pattern (56 of 233 ethanol-preserved samples, 24.0%; 16 of 187 dry-preserved samples, 8.6%) that was also not statistically significant (age: likelihood-ratio statistic = 0.911, $p = 0.3398$; preservation: likelihood-ratio statistic = 1.117, $p = 0.5720$). Human and *Cutibacterium* detections were, however, strongly associated with one another (Fisher's exact test, odds ratio = 28.47, 95% CI = 14.69–56.96, $p < 2.2 \times 10^{-16}$), consistent with *Cutibacterium* representing a human skin-associated contaminant (Table S11).

Contaminant (non-octocoral) *COI* barcodes were recovered from 4.2% of samples (5.1% of genome-skimmed samples, 0.8% of target-captured samples; Figure 4c), and were uncommon overall (Table S8). Most *COI* contaminants (70%) were identified as Symbiodiniaceae (all from zooxanthellate species), and 25% were non-coral eukaryotes (e.g., sponges, multicellular algae) that could have been growing in association with the coral. Contaminant *28S rDNA* barcode sequences were recovered from 15.9% of samples, but in a much higher fraction of genome-skimmed (18.9%) samples compared to those that had been target-captured (4.9%) (Figure 4d). Recovery of contaminant *28S* sequences increased in the raw age-stratified frequencies, reaching 27% among specimens collected > 100 years ago (Figure 7a), but this age effect was not retained strongly in the final adjusted model (binomial GLMM, likelihood-ratio statistic = 1.968, $p = 0.1607$; Table S8). *28S* contaminants included fungal (29.5%), human (27.3%), and Symbiodiniaceae (22.7%) sequences. Fungal *28S* sequences were more often recovered from dry-preserved material, while human sequences were more often recovered from ethanol-preserved material. The frequency and identity of contaminant *28S* sequences also varied among museum repositories (Figure 7b), with > 25% of samples from BMNH, MCZ, MZUT, and ZMB exhibiting contamination. These patterns should, however, be interpreted cautiously because some categories were sparse, and adjusted museum effects on *28S* contamination were weak (Table S7). Only fungal contaminants were recovered from MZUT, while fungal, human, and other eukaryote 28S sequences were found in specimens from the other three repositories.

3.5 | Phylogenetic Placement

A majority of the samples with > 100 UCE loci ($n = 362$) were recovered in the expected clade in maximum likelihood analyses that included just those samples with a published reference dataset (Quattrini et al. 2024; Figure S3). Several specimens that were highly supported in unexpected phylogenetic positions were subsequently confirmed by morphological re-examination to have been misclassified (e.g., assigned to the incorrect genus) by the original taxonomic authority (see Discussion). Twenty-three (7%) of the samples with > 100 UCE loci were, however, recovered in a clade that was sister to all other octocorals. BLAST analyses confirmed that a large fraction (55.1%–95.5%) of the UCE loci recovered from these 23 samples were contaminant sequences that did not match octocoral UCEs (Table S12). Only 4 to 45% (median 9%) of loci matched coral sequences. The remaining loci BLASTed with high similarity (> 70% identity) to human (1%–92%, median 43%) or to prokaryotic and other non-coral genomes (2%–94%, median 42%). When maximum likelihood analysis was repeated following the removal of the non-octocoral UCE loci from these 23 samples, eight (35%) were

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

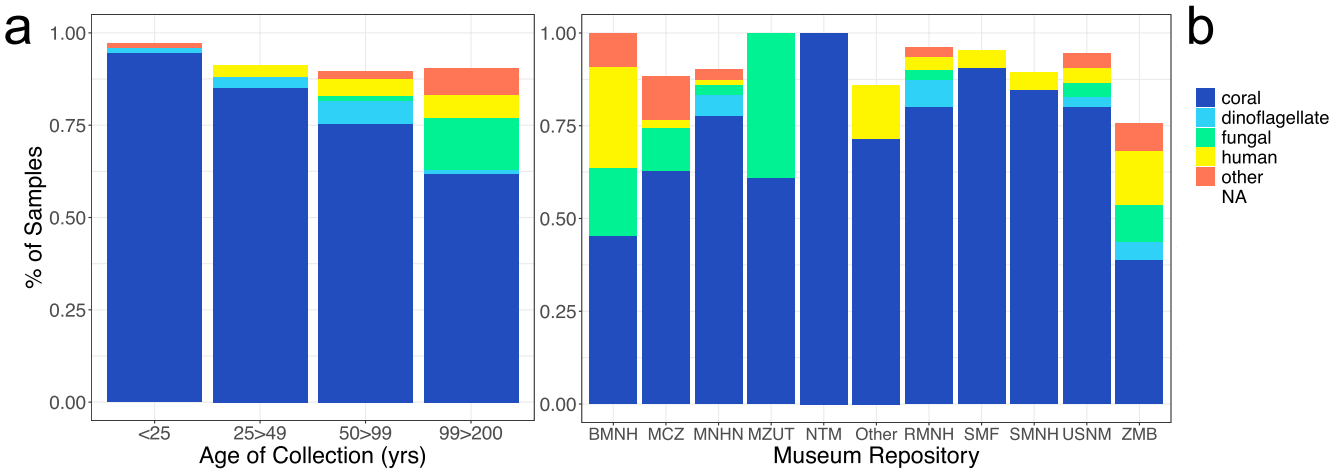


FIGURE 7 | Identity of *28S rDNA* barcode sequences recovered from octocoral type specimens (a) of different collection ages and (b) from different museum repositories. Samples from which no *28S rDNA* was recovered (NA) are included in total.

recovered in the expected clade in the phylogeny (although often on anomalously long branches; Figure S4). Among the remaining 15 samples (65%), six formed a spurious clade within Malacalcyonacea while the other nine were distributed throughout the phylogeny in incorrect clades. Maximum likelihood analyses that included all samples with > 5 UCE loci or only those with 5–100 UCE loci failed to converge. When trees were subset to include only samples with 5–100 UCE loci plus reference taxa from specific clades (e.g., Sarcophytidae, Acanthogorgiidae), a majority of the types were still recovered in spurious clades sister to the reference clade.

Among the 362 samples with > 100 UCE loci, *mtMutS* and *28S rDNA* barcodes identified (by BLAST) as octocoral were recovered from 280 (77%) and 320 (88%), respectively, including 11 of the 23 samples that had been placed in a spurious sister clade in the UCE phylogeny (Table S12). All of these *mtMutS* sequences and all but two *28S* sequences were placed in the expected clades in maximum likelihood trees (Figures S5 and S6); in six cases, phylogenetic position was inferred to be correct only after re-examination of the specimen and reassessment of its generic assignment. Among the 181 samples from which we recovered < 100 UCE loci, *mtMutS* barcodes were recovered from 43 (24%) and *28S rDNA* from 83 (46%). All but one *mtMutS* and three *28S* sequences were placed in the expected clade in the maximum likelihood trees.

4 | Discussion

We demonstrate the feasibility and relative ease of obtaining genomic data from historic, ethanol- and dry-preserved type material of octocorals. Genomic characterization of name-bearing types amplifies the value of these exemplar specimens and provides a novel and important character set needed for reliable species identification. The addition of sequence data from type specimens to data repositories such as NCBI and BOLD (as well as to museum databases) will increase the accuracy of barcode-based species identifications and help to flag misidentified sequences that are already in those repositories (Renner et al. 2024). Moreover, genomic data captured from types provide taxonomists with a set of molecular characters that will facilitate assignment of correct binomials to species that have proven difficult to discriminate using morphology alone (e.g., McFadden et al. 2017).

4.1 | Factors Influencing Recovery of Genomic Data From Historical Type Material

Method of preservation was more important than the age of the collection as a predictor of the amount and quality of genomic data that could be recovered from types, with overall higher recovery rates and no age-related decline in recovery of UCE loci, mitogenomes and barcode sequences from dry-preserved specimens. While preserving octocoral specimens dry was a frequent practice in the 19th and early 20th centuries, in recent decades most type material has instead been preserved in 70%–95% ethanol (Figure 2). Fluid media undoubtedly better preserve the integrity of soft tissues and, in colonies that lack a rigid skeletal axis, the overall form of the colony, both of which may be distorted by desiccation. The calcitic sclerites that are the primary morphological characters used for taxonomic classification and species-level identification of octocorals are not, however, damaged by drying. In light of our results, we suggest that subsamples of type material should also be preserved dry for long-term preservation of DNA, particularly when access to dedicated biorepositories is unavailable, with additional subsamples preserved in ethanol to maintain the structure of soft tissues such as polyps.

Historically, formalin has not been widely used for fixation of octocoral tissues due to its potential to acidify over time, thereby damaging the calcitic sclerites (Fabricius and Alderslade 2001), and, more recently, recognition of its negative effects on DNA (Srinivasan et al. 2002). Although we did recover UCEs, PCGs and barcode loci from material known to have been fixed in formalin, recovery from that subset of samples was generally poor and loci were typically shorter. These patterns should be interpreted cautiously, however, because few specimens ($n = 15$) could be confidently identified as formalin-fixed. The preservation history of older material

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

housed in museum collections is rarely documented and it is likely that some of the ethanol-preserved specimens we sequenced may have been fixed in formalin originally. Likewise, whether dry specimens were originally preserved in ethanol or fixed in formalin and subsequently allowed to dry out usually cannot be determined from the metadata associated with the material. We observed a very wide range of outcomes (e.g., numbers of UCEs and PCGs recovered) from ethanol-preserved specimens collected 50–75 years ago (Figure 5) that could be indicative of some of that material having been formalin-fixed prior to storage in ethanol.

Our results suggest that genome-skimming (i.e., low-coverage shotgun sequencing) was a more effective method than target capture for sequencing historical type material of octocorals under the protocols used here, in particular for material > 50 years old. Target capture preferentially amplifies the targeted (i.e., UCE) loci prior to sequencing, reducing the number of off-target sequence reads recovered and increasing sequencing depth of the targeted loci (McCormack et al. 2012). While this approach may reduce the number of contaminant sequences recovered (Letch et al. 2025), it also reduces the numbers of reads obtained for mitochondrial or rDNA genes that have not been targeted explicitly. Genome-skimming thus offers a significant advantage over target capture by more consistently recovering mt genomes and barcodes that can be used to include type material in phylogenetic and species delimitation analyses unless the target-capture bait sets incorporate mt and rDNA genes. These conventionally used barcodes are important for populating reference libraries that are necessary for environmental (eDNA) and other metabarcoding approaches (Everett and Park 2018; McCartin et al. 2024, 2026). Moreover, while target capture and genome-skimming both recovered high numbers of UCE loci in specimens collected relatively recently (< 50 years ago), target capture performed relatively poorly with much older specimens, recovering fewer and shorter UCE loci in material collected > 50 years ago. In addition to differences in data yield, genome-skimming and target capture differ in cost and processing time. Library preparation for genome-skimming is generally simpler, faster, and less expensive per sample than target capture, which requires additional enrichment steps and specialized bait reagents (Quattrini et al. 2018, 2024). Our results support the conclusions of Letch et al. (2025) that genome-skimming represents a more cost-effective first step to sequencing type material. Skimming enables the recovery of more mitogenomes and/or barcode data per sample and (in material > 50 years old) substantially more UCE loci than target capture.

This comparison should not be interpreted as indicating that target capture is unsuitable for degraded museum material under all circumstances. The target-capture protocol used here was not specifically optimized for highly fragmented historical DNA. Recovery from older specimens could potentially be improved by using bait sets designed for shorter templates, increasing bait tiling density, modifying hybridization stringency and duration, increasing probe concentration, using library-preparation protocols tailored to low-input degraded DNA, optimizing blocking oligos and PCR conditions to reduce off-target and duplicate reads, and not pooling samples for target enrichment. Therefore, our comparison reflects the relative performance of genome-skimming and target capture as implemented in this study, rather than an absolute limit on the performance of target capture for historical material of octocorals or other taxa (Connelly et al. 2024).

Corals are holobionts that form symbioses with both prokaryotes and other eukaryotes, and DNA isolated from the host naturally includes the genomes of those associated organisms. Among the zooxanthellate species we sequenced (i.e., those forming photosymbioses with dinoflagellates), recovery of sequence reads from Symbiodiniaceae was high, representing > 80% of reads in some samples (Figure 6). Rather than being an impediment to phylogenomics, these off-target reads represent valuable additional information, allowing the identity of Symbiodiniaceae associated with the host octocoral to be characterized in most cases (e.g., Connelly et al. 2026; Table S9). We recovered *Cladocopium* sequences from most of the Indo-Pacific zooxanthellate species along with *Breviolum*, *Durusdinium* and *Symbiodinium* in some species from either the Pacific or tropical western Atlantic, reflecting the known associations of these photosymbiont genera with the octocoral taxa included in this study (Goulet et al. 2008). Furthermore, other naturally occurring symbionts were detected including corallicolid apicomplexans, Endozoicomonadaceae, and Mollicutes by assembling their SSU rRNA genes (Table S10). Such ancillary information is extremely valuable for characterization of the entire holobiont, knowledge that is necessary to predict and mitigate effects of environmental change on coral communities (Peixoto et al. 2017; Voolstra and Ziegler 2020).

Some of the off-target sequences we detected and assembled were known contaminants and not likely to be naturally occurring associates including human DNA, human-associated bacteria (*Cutibacterium acnes* and *Staphylococcus*), and fungus (*Aspergillus*, *Phialosimplex* [=*Aspergillus*], *Malassezia*). While reagents and DNA can be contaminated during the sequencing process, the long-term storage and handling of specimens over time likely contributes more to contamination of types by fungal and human DNA (Chaudary et al. 2019; Keles and Celik 2026; Tulloch et al. 2026). We encourage museum scientific staff and visitors to consistently use gloves when handling specimens. While specimens preserved in ethanol yielded poorer quality coral DNA and were more prone to contamination by the DNA of humans and human-associated microbes, they were nearly free of Aspergillaceae which were far more likely to grow on dry specimens where they were probably contributing to their degradation (Romero et al. 2021). These are valuable insights to consider both for interpreting museomic sequence data and for developing long-term conservation strategies for historical type material.

4.2 | Phylogenetic Insights Gained by Sequencing Type Material

The amount and quality of genomic data that can be obtained from historical specimens impacts their use for phylogenetic inference. Recovery of hundreds of nuclear loci plus mitochondrial protein coding genes allowed types to be included in phylogenomic analyses and typically provided strong support for their placement within those phylogenies, leading to novel insights on the identity of enigmatic and misclassified species. Taxa

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from which relatively few (e.g., < 100) UCE loci were recovered were, however, often placed in spurious positions in phylogenomic analyses, either on very long branches or in unique clades where they were grouped with other samples that had few loci. These spurious placements are likely due to some combination of poor sequence quality of the loci that were recovered leading to poor sequence alignments and to recovery of UCE loci from non-target organisms (e.g., other members of the holobiont or contaminants).

Further investigation into the spurious placement of 23 samples with > 100 UCEs that formed a clade sister to all other octocorals in the phylogeny revealed that a majority of their UCE loci did not match coral genomes. These results suggest that the Phyluce pipeline retained substantial numbers of non-coral UCE loci despite filtering thresholds of > 70% identity and > 70% coverage. Across samples, many loci appeared to correspond to conserved or shared genomic regions from off-target taxa often associated with museum specimens, including human and microbial sources (Table S12). We therefore urge caution when interpreting phylogenetic placements inferred from datasets that include high proportions of contaminant loci. These results do not diminish the value of sequence data obtained from type material, but rather emphasize that sequences derived from types should be interpreted together with the morphology, labels and collection history of the corresponding name-bearing specimens, particularly when locus recovery is low, contamination is detected, or phylogenetic placement is unexpected.

Some of the samples from which few UCE loci or high proportions of contaminant loci were recovered may nonetheless still retain phylogenetically informative data, and it is possible to salvage the information recovered from these samples in other ways. For datasets with substantial contamination, UCE loci could be screened against available coral genomes and only those identified as coral retained for downstream phylogenetic analyses, as shown in this study. Additional improvements to sequence quality may be possible for preserved museum specimens, which often have cytosine deamination sequencing errors (Briggs et al. 2007). Sequence data could be evaluated using MapDamage (Jónsson et al. 2013) and corrected through damage-aware rescaling or end trimming, following approaches commonly applied in museomic studies (e.g., Kistler et al. 2025). The corrected reads could then be remapped to UCE reference loci using pipelines such as Phyluce (Faircloth et al. 2012). Although these approaches were not evaluated herein, they may provide opportunities to recover additional loci from challenging specimens and improve phylogenetic placement.

In cases where too few UCE loci are recovered for reliable phylogenetic inference, recovery of barcode loci from types can nonetheless provide valuable information about a species' identity and phylogenetic placement. In almost all cases for which we obtained either *mtMutS* or *28S rDNA* barcodes for a specimen with < 100 UCE loci—most of which were placed in spurious positions in the UCE tree—inclusion of the barcode sequence in a gene tree for the locus recovered the type specimen in the expected clade. In addition to helping to resolve phylogenetic affinities, single-locus barcodes recovered from type material also provide invaluable resources for species identification, necessary to validate the identities of reference sequences in databases that are used to assess biodiversity via eDNA and other metabarcoding approaches (Renner et al. 2024).

In a few cases we recovered UCE or barcode sequences that strongly supported a type specimen in an unexpected position in the phylogeny. In some cases, the unexpected outcome could be traced to a lab error (e.g., mislabeling or switching of samples), in which case data were removed from all analyses and the specimen targeted for re-sequencing. In other cases, however, re-examination of a specimen revealed that its original genus assignment had been erroneous and the unexpected phylogenetic position was supported a posteriori by morphology. For example, a syntype of *Eunicea ransoni* Stiasny, 1937 was recovered in the genus *Muricea*, consistent with a handwritten note of unknown provenance that had been found in the jar identifying it as *Muricea appressa*. Similarly, two sequenced syntypes of *Thesea guadalupensis* Duchassaing & Michelotti, 1864 were not recovered as sister taxa. When reviewing the literature, we discovered that Deichmann (1936) had examined type material and commented that a fragment housed at MCZ was *T. guadalupensis*, whereas the specimen in the British Museum originating from Turin and labelled *T. guadalupensis* was in fact *Acanthacis scabra* (Deichmann 1936: 120, 131). Subsequent morphological examination has confirmed the MZUT specimen to be *Acanthacis* (R. Cordeiro, pers. com.), and this species was in fact recovered in the expected family, Acanthogorgiidae, in the *mtMutS* phylogeny.

In several cases, the phylogenetic affinities of species considered to be *taxa inquirenda* (names of uncertain application) were suggested by a barcode or UCEs and confirmed following reassessment of morphology. These included the verification that *Ammothea parasitica* Duchassaing & Michelotti, 1860 is a species of *Briareum* and recognition that a syntype specimen of *Plexaura michelini* Stiasny, 1937 is actually the skeleton of a gorgonian that had been overgrown by the epifaunal soft coral *Alcyonium coralloides*. These and many other historical types are still in need of redescription using modern imaging technologies (e.g., scanning electron microscopy (SEM), micro-CT, photogrammetry) to confirm, update and widen the suite of morphological characters available for species identification.

Coauthors and collaborators of the current study are actively generating SEM images of the sequenced type material to make publicly available in museum databases (e.g., Lazar et al. 2025), while some species, genera, and families are being redescribed, including *Klyxum*, *Paramuricea, Placogorgia*, and *Thesea*. The morphological redescriptions necessary for validation and formal revisions of taxa require considerably more time and taxon-specific expertise than genome sequencing, however, making the genomic data an invaluable resource for identifying species while documentation of their morphology proceeds more slowly.

5 | Conclusions

We have demonstrated that historical type specimens of octocorals can yield substantial genomic data, including hundreds of UCE loci, mitochondrial protein-coding genes, complete mitogenomes, and DNA barcoding loci, even when specimens were collected centuries ago. The quantity and quality of genomic data recovered depend more strongly on preservation

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

method and sequencing approach than on specimen age alone. Dry-preserved material showed the least age-related decline in recovery of genomic data, while genome-skimming outperformed target-capture for recovery of UCE loci in older specimens. Furthermore, genome-skimming can recover partial to full mitogenomes and/or *28S rDNA*, all important barcodes for species identification and eDNA analyses, particularly when UCE loci recovery fails. These findings demonstrate that museum type collections are archives of both morphological data and invaluable genomic resources that can be used to stabilize names. More broadly, the approaches developed here provide a practical framework for unlocking genomic data from historical collections of other soft-bodied marine invertebrates.

Author Contributions

C.S.M., A.M.Q. and Y.B. conceived the study; C.S.M., D.M., C.F.V., T.C., K.A.C., E.S.G., L.M., Y.B., M.S. and H.H.W. conducted laboratory work; C.S.M., A.M.Q., D.M., K.S.-N. and S.A.V. analysed data; C.S.M., A.M.Q., D.M., and K.S.-N. wrote the manuscript with contributions from all authors. All authors have read, revised and approved the submitted manuscript.

Acknowledgements

We thank the many museum curators and collection managers who have assisted and graciously allowed us to sample tissues in their type collections, in particular Z. Kuplik (SMNHTAU); M. Castelin (MNHN); B. Hoeksema and H. Bakker (RMNH); C. Lüter and A. Schwiering (ZMB); M. Lowe (BMNH); H. Sattmann and P. Frade (NHMW); A. Schmidt-Rhaesa and H. Roggenbuck (ZMH); B. Ruthensteiner (ZSM); M. V. Sørensen and L. Pavesi (ZMUC); M. Ekins (QM); P. Alderslade and S. Horner (NTM); R. Tota, S. Scapinello and F. Andreone (MZUT); D. Janussen, N. Santodomingo and S. Dimter (SMF); M. Bruni (Oceanographic Institute Foundation Albert I of Monaco); W. Moser, H. Shull, A. Robinson, and S. Bryant (USNM); and A. Baldinger and the MCZ Department of Invertebrate Zoology and MCZ Cryogenic collection. NOAA Office of Education Educational Partnership Program award (NA21SEC4810004) supported E.S.G. K.S.-N. thanks the Richard Lounsbery Foundation for support and, in this regard, acknowledges J. H. Ausubel (Rockefeller University) and L. Brown (Lounsbery Foundation) for their encouragement. Molecular work was undertaken in part at the Laboratories of Analytical Biology (LAB) at the NMNH. The computations in this paper were conducted on the Smithsonian High-Performance Cluster (SI/HPC), Smithsonian Institution (see https://doi.org/10.25572/SIHPC). Sequencing was procured and managed via Genohub.com. We thank the editor and the anonymous reviewers for constructive comments that helped improve the manuscript.

Funding

Molecular work and analyses were funded in part by U.S. National Science Foundation DEB-1929319, U.S.-Israeli Binational Science Foundation 2019624, the Smithsonian NMNH Department of Invertebrate Zoology Coral Endowment Fund and the U.S. Department of the Interior, Bureau of Ocean Energy Management (BOEM), Environmental Studies Program, Washington, DC, through Interagency Agreement Number M20PG00008 with the National Museum of Natural History Department of Invertebrate Zoology (NMNH IZ). T.C. received funding from the Regional Government of the Azores (ACORES-14-62G4-FEDER-000001; 18/DRPM/2022) and the FAM Foundation (2024/25). Okeanos Institute received funds through the Portuguese Foundation for Science and Technology, project reference UID/05634/2025 (DOI: 10.54499/UID/05634/2025), and through the Azores Regional Directorate for Science, Innovation and Development via the PROSCIENTIA Incentive System, project M1.1.A/FUNC.UI&D/014/2025. L.M. travel and sequencing costs were supported by KAUST baseline research funds BAS/1/1090-01-01 to F. Benzoni. K.S.-N. was partially funded by a grant from the Richard Lounsbery Foundation, as well as a *Chercheur invite* appointment from the Department Origine & Évolution at MNHN, Paris (AAP fédérateur DGD-REV 2024, PI: M. Castelin), and a "projet d'établissement" by Sorbonne University. Funding to E.S.G. (through E. Easton) was by the National Oceanic and Atmospheric Administration, Office of Education Educational Partnership Program award (NA21SEC4810004). K.A.C. was funded in part by NOAA's National Centers for Coastal Ocean Science, Competitive Research Program and Office of Ocean Exploration and Research award (NA18NOS4780166, S. Herrera PI). This work was also funded by King Abdullah University of Science and Technology (grant BAS/1/1090-01-01).

Ethics Statement

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Raw sequence reads are deposited in the SRA under BioProjects PRJNA1466401, PRJNA1467188, PRJNA1460859, PRJNA1365937, PRJNA822352, and PRJNA588468. Assembled mitogenomes and barcode loci are available in GenBank (see Table S1). Sequence alignments, treefiles, supplementary statistical tables, and code used for all analyses are available at dryad, DOI: 10.5061/dryad.p5hqbzm3p.

Benefits from this research accrue from the sharing of our data and results in public databases as described above.

References

Ahrens, D., E. Bazzato, A. C. Castro Lopez, M. Krause, C. Ancona, and L. Dietz. 2026. "Integrative Taxonomy of Cryptic *Pachypus* Chafers Using Museomics, Morphometrics, Barcoding, and Genomic DNA Analysis (Coleoptera: Scarabaeidae: Pachypodinae)." *Scientific Reports* 16: 15710.

Alderslade, P. 2001. "Six New Genera and Six New Species of Soft Coral, and Some Proposed Familial and Subfamilial Changes Within the Alcyonacea (Coelenterata: Octocorallia)." *Bulletin of the Biological Society of Washington* 10: 15–65.

Allio, R., A. Schomaker-Bastos, J. Romiguier, F. Prosdocimi, B. Nabholz, and F. Delsuc. 2020. "MitoFinder: Efficient Automated Large-Scale Extraction of Mitogenomic Data in Target Enrichment Phylogenomics." *Molecular Ecology Resources* 20, no. 4: 892–905. https://doi.org/10.1111/1755-0998.13160.

Bankevich, A., S. Nurk, S. Antipov, et al. 2012. "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing." *Journal of Computational Biology* 19: 455–477.

Benayahu, Y., C. S. McFadden, and E. Shoham. 2017. "Search for Mesophotic Octocorals (Cnidaria, Anthozoa). I. A New Sclerite-Free Genus From Eilat, Northern Red Sea." *ZooKeys* 680: 1–11.

Benayahu, Y., L. P. van Ofwegen, and C. S. McFadden. 2018. "Evaluating the Genus *Cespitularia* and Descriptions of New Genera of the Family Xeniidae (Octocorallia, Alcyonacea)." *ZooKeys* 754: 63–101.

Bickford, D., D. J. Lohman, N. S. Sodhi, et al. 2007. "Cryptic Species as a Window on Diversity and Conservation." *Trends in Ecology & Evolution* 22, no. 3: 148–155.

Bridge, T. C. L., P. F. Cowman, A. M. Quattrini, et al. 2024. "A *tenuis* Relationship: Traditional Taxonomy Obscures Systematics and Biogeography of the '*Acropora tenuis*' (Scleractinia: Acroporidae)

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Species Complex." *Zoological Journal of the Linnean Society* 202, no. 2: zlad062. https://doi.org/10.1093/zoolinnean/zlad062.

Briggs, A. W., U. Stenzel, P. L. F. Johnson, and S. Pääbo. 2007. "Patterns of Damage in Genomic DNA Sequences From a Neanderthal." *Proceedings of the National Academy of Sciences of the United States of America* 104, no. 37: 14616–14621.

Call, E. 2020. "The Age of Museomics: How to Get Genomic Information From Museum Specimens of Lepidoptera." Ph.D. Thesis, Department of Biology, Lund University. https://portal.research.lu.se/en/publications/the-age-of-museomics-how-to-get-genomic-information-from-museum-s/.

Camacho, C., G. Coulouris, V. Avagyan, et al. 2009. "BLAST+: Architecture and Applications." *BMC Bioinformatics* 10: 421. https://doi.org/10.1186/1471-2105-10-421.

Carlo, J. M., M. S. Barbeitos, and H. R. Lasker. 2011. "Quantifying Complex Shapes: Elliptical Fourier Analysis of Octocoral Sclerites." *Biological Bulletin* 220: 224–237. https://doi.org/10.1086/BBLv220n3p224.

Castañeda-Rico, S., C. W. Edwards, M. T. R. Hawkins, and J. E. Maldonado. 2022. "Museomics and the Holotype of a Critically Endangered Cricetid Rodent Provide Key Evidence of an Undescribed Genus." *Frontiers in Ecology and Evolution* 10: 930356. https://doi.org/10.3389/fevo.2022.930356.

Castresana, J. 2000. "Selection of Conserved Blocks From Multiple Alignments for Their Use in Phylogenetic Analysis." *Molecular Biology and Evolution* 17: 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334.

Chaudary, R., I. Gautam, R. Tuladhar, et al. 2019. "Microbiota Diversity on the Preserved Specimens of Natural History Museum, Swayambhu, Kathmandu, Nepal." *International Journal of Pharma and Bio Sciences* 9, no. 2: 51–61. https://doi.org/10.21276/ijpbs.2019.9.sp2.6.

Chen, S., Y. Zhou, Y. Chen, and J. Gu. 2018. "Fastp: An Ultra-Fast All-In-One FASTQ Preprocessor." *Bioinformatics* 34, no. 17: i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Coates, D. J., M. Byrne, and C. Moritz. 2018. "Genetic Diversity and Conservation Units: Dealing With the Species-Population Continuum in the Age of Genomics." *Frontiers in Ecology and Evolution* 6: 165. https://doi.org/10.3389/fevo.2018.00165.

Connelly, M. T., M. G. Catapang, and A. M. Quattrini. 2024. "Unlocking the Treasure Trove: Leveraging Dry Coral Specimens for Museum Genomics." *Coral Reefs* 43, no. 4: 1153–1159. https://doi.org/10.1007/s00338-024-02525-5.

Connelly, M. T., V. M. Glynn, A. Cornejo, et al. 2026. "Genome Skimming Illuminates Hidden Species Diversity and Symbiodiniaceae Associations in East Pacific *Pocillopora* Corals." *Genome Biology and Evolution* 18, no. 2: evaf235. https://doi.org/10.1093/gbe/evaf235.

Costa, H., G. M. Foody, S. Jiménez, and L. Silva. 2015. "Impacts of Species Misidentification on Species Distribution Modeling With Presence-Only Data." *ISPRS International Journal of Geo-Information* 4, no. 4: 2496–2518. https://doi.org/10.3390/ijgi4042496.

de Queiroz, K. 2007. "Species Concepts and Species Delimitation." *Systematic Biology* 56: 879–886. https://doi.org/10.1080/10635150701701083.

Deichmann, E. 1936. "The Alcyonaria of the Western Part of the Atlantic Ocean." *Memoirs of the Museum of Comparative Zoology, Harvard College* 53: 1–317.

Erickson, K. L., A. Pentico, A. M. Quattrini, and C. S. McFadden. 2021. "New Approaches to Species Delimitation and Population Structure of Anthozoans: Two Case Studies of Octocorals Using Ultraconserved Elements and Exons." *Molecular Ecology Resources* 21: 78–92. https://doi.org/10.1111/1755-0998.13241.

Everett, M. V., and L. K. Park. 2018. "Exploring Deep-Water Coral Communities Using Environmental DNA." *Deep-Sea Research Part II: Topical Studies in Oceanography* 150: 229–241. https://doi.org/10.1016/j.dsr2.2017.09.008.

Fabricius, K., and P. Alderslade. 2001. *Soft Corals and Sea Fans: A Comprehensive Guide to the Tropical Shallow-Water Genera of the Central-West Pacific, the Indian Ocean and the Red Sea*, 264. Australian Institute of Marine Science (AIMS).

Faircloth, B. C. 2016. "PHYLUCE Is a Software Package for the Analysis of Conserved Genomic Loci." *Bioinformatics* 32: 786–788. https://doi.org/10.1093/bioinformatics/btv646.

Faircloth, B. C., J. E. McCormack, N. G. Crawford, M. G. Harvey, R. T. Brumfield, and T. C. Glenn. 2012. "Ultraconserved Elements Anchor Thousands of Genetic Markers Spanning Multiple Evolutionary Timescales." *Systematic Biology* 61, no. 5: 717–726.

Glenn, T. C., R. A. Nilson, T. J. Kieran, et al. 2019. "Adapterama I: Universal Stubs and Primers for 384 Unique Dual-Indexed or 147,456 Combinatorially-Indexed Illumina Libraries (iTru & iNext)." *PeerJ* 7: e7755. https://doi.org/10.7717/peerj.7755.

Goulet, T. L., C. Simmons, and D. Goulet. 2008. "Worldwide Biogeography of *Symbiodinium* in Tropical Octocorals." *Marine Ecology Progress Series* 355: 45–58. https://doi.org/10.3354/meps07214.

Gruber-Vodicka, H. R., B. K. B. Seah, and E. Pruesse. 2020. "phyloFlash: Rapid SSU rRNA Profiling and Targeted Assembly From Metagenomes." *mSystems* 5: e00920-20. https://doi.org/10.1128/mSystems.00920-20.

Hahn, C., L. Bachmann, and B. Chevreux. 2013. "Reconstructing Mitochondrial Genomes Directly From Genomic Next-Generation Sequencing Reads—A Baiting and Iterative Mapping Approach." *Nucleic Acids Research* 41, no. 13: e129. https://doi.org/10.1093/nar/gkt371.

Hahn, E. E., M. R. Alexander, A. Grealy, J. Stiller, D. M. Gardiner, and C. E. Holleley. 2021. "Unlocking Inaccessible Historical Genomes Preserved in Formalin." *Molecular Ecology Resources* 22: 2130–2147. https://doi.org/10.1111/1755-0998.13505.

Hoang, D. T., O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh. 2018. "UFBoot2: Improving the Ultrafast Bootstrap Approximation." *Molecular Biology and Evolution* 35: 518–522. https://doi.org/10.1093/molbev/msx281.

ICZN. 1999. *International Code of Zoological Nomenclature*. 4th ed. International Trust for Zoological Nomenclature.

Jónsson, H., A. Ginolhac, M. Schubert, P. L. F. Johnson, and L. Orlando. 2013. "mapDamage2.0: Fast Approximate Bayesian Estimates of Ancient DNA Damage Parameters." *Bioinformatics* 29: 1682–1684. https://doi.org/10.1093/bioinformatics/btt193.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. "ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates." *Nature Methods* 14: 587–589. https://doi.org/10.1038/nmeth.4285.

Katoh, K., and D. M. Standley. 2013. "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability." *Molecular Biology and Evolution* 30: 772–780.

Keles, E., and O. Celik. 2026. "Metagenomic and Microbiological Analyses of Historical Manuscripts for Bacterial Community Profiling and Bacteria-Related Biodeterioration Assessment." *Microbial Cell* 13: 117–130. https://doi.org/10.15698/mic2026.03.871.

Kistler, L., F. O. Freitas, R. M. Gutaker, et al. 2025. "Historic Manioc Genomes Illuminate Maintenance of Diversity Under Long-Lived Clonal Cultivation." *Science* 387, no. 6738: eadq0018.

Lazar, E., C. S. McFadden, D. Huchon, and Y. Benayahu. 2025. "Redescription of Type Material of the Genus *Cladiella* Gray, 1869

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Octocorallia, Malacalcyonacea, Cladiellidae) From the Red Sea." *Zootaxa* 5674: 1–84. https://doi.org/10.11646/zootaxa.5674.1.1.

Letch, H., C. Greve, A. K. Hundsdoerfer, I. Irisarri, J. M. Moore, and S. U. Pauls. 2025. "Type Genomics: A Framework for Integrating Genomic Data Into Biodiversity and Taxonomic Research." *Systematic Biology* 74: 1029–1044.

Linnaeus, C. 1758. *Systema Naturae Per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, Cum Characteribus, Differentiis, Synonymis, Locis*, 824. Laurentii Salvii.

McCartin, L., E. Saso, S. Vohsen, et al. 2024. "Nuclear eDNA Metabarcoding Primers for Anthozoan Coral Biodiversity Assessment." *PeerJ* 12: e18607. https://doi.org/10.7717/peerj.18607.

McCartin, L. J., S. A. Vohsen, A. L. Wood, et al. 2026. "Accounting for Intra- and Intergenomic Sequence Variation in Reference Barcodes Improves eDNA Metabarcoding Biodiversity Assessment." *Molecular Ecology Resources* 26: e70130. https://doi.org/10.1111/1755-0998.70130.

McCormack, J. E., S. M. Hird, A. J. Zellmer, B. C. Carstens, and R. T. Brumfield. 2012. "Applications of Next-Generation Sequencing to Phylogeography and Phylogenetics." *Molecular Phylogenetics and Evolution* 66: 526–538. https://doi.org/10.1016/j.ympev.2011.12.007.

McFadden, C. S., Y. Benayahu, E. Pante, J. N. Thoma, P. A. Nevarez, and S. C. France. 2011. "Limitations of Mitochondrial Gene Barcoding in Octocorallia." *Molecular Ecology Resources* 11: 19–31.

McFadden, C. S., R. Cordeiro, K. Samimi-Namin, B. W. Hoeksema, G. Williams, and L. P. van Ofwegen. 2026. "World List of Octocorallia." https://www.marinespecies.org/octocorallia.

McFadden, C. S., R. Haverkort-Yeh, A. M. Reynolds, et al. 2017. "Species Boundaries in the Absence of Morphological, Ecological or Geographical Differentiation in the Red Sea Octocoral Genus *Ovabunda* (Alcyonacea: Xeniidae)." *Molecular Phylogenetics and Evolution* 112: 174–184. https://doi.org/10.1016/j.ympev.2017.04.025.

McFadden, C. S., and L. P. van Ofwegen. 2013. "A Second, Cryptic Species of the Soft Coral Genus *Incrustatus* (Anthozoa: Octocorallia: Clavulariidae) From Tierra del Fuego, Argentina Revealed by DNA Barcoding." *Helgoland Marine Research* 67: 137–147.

McFadden, C. S., L. P. van Ofwegen, E. J. Beckman, Y. Benayahu, and P. Alderslade. 2009. "Molecular Systematics of the Speciose Indo-Pacific Soft Coral Genus, *Sinularia* (Anthozoa: Octocorallia)." *Invertebrate Biology* 128: 303–323. https://doi.org/10.1111/j.1744-7410.2009.00179.x.

McFadden, C. S., L. P. van Ofwegen, and A. M. Quattrini. 2022. "Revisionary Systematics of Octocorallia (Cnidaria: Anthozoa) Guided by Phylogenomics." *Bulletin of the Society of Systematic Biologists* 1, no. 4: 8735. https://doi.org/10.18061/bssb.v1i3.8735.

Minh, B. Q., H. A. Schmidt, O. Chernomor, et al. 2020. "IQTree-2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era." *Molecular Biology and Evolution* 37: 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Morrissey, D., C. F. Vaga, K. A. Coil, H. Wirshing, and A. M. Quattrini. 2026. "Improved Molecular Kit-Based Workflow to Unlock the Genetic Potential of Historical Coral Type Specimens." *BMC Research Notes* 19: 271.

Palandačić, A., M. J. Chai, G. A. Shandikov, et al. 2024. "An Annotated Catalog of Selected Historical Type Specimens, Including Genetic Data, Housed in the Natural History Museum Vienna." *ZooKeys* 1203: 253–323.

Peixoto, R. S., P. M. Rosado, D. C. A. Leite, A. S. Rosado, and D. G. Bourne. 2017. "Beneficial Microorganisms for Corals (BMCs): Proposed Mechanisms for Coral Health and Resilience." *Frontiers in Microbiology* 8: 341.

Quattrini, A. M., B. Faircloth, L. Dueñas, et al. 2018. "Universal Target-Enrichment Baits for Anthozoan (Cnidaria) Phylogenomics: New Approaches to Long-Standing Problems." *Molecular Ecology Resources* 18: 281–295.

Quattrini, A. M., L. J. McCartin, E. E. Easton, et al. 2024. "Skimming Genomes for Systematics and DNA Barcodes of Corals." *Ecology and Evolution* 14: e11254. https://doi.org/10.1002/ece3.11254.

R Core Team. 2026. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Raxworthy, C. J., and B. T. Smith. 2021. "Mining Museums for Historical DNA: Advances and Challenges in Museomics." *Trends in Ecology & Evolution* 36: 1049–1060. https://doi.org/10.1016/j.tree.2021.07.009.

Renner, S. S., M. D. Scherz, C. L. Schoch, M. Gottschling, and M. Vences. 2024. "Improving the Gold Standard in NCBI GenBank and Related Databases: DNA Sequences From Type Specimens and Type Strains." *Systematic Biology* 73: 486–494. https://doi.org/10.1093/sysbio/syad068.

Rodrigues, A. V., G. Nakamura, V. G. Staggemeier, and L. Duarte. 2022. "Species Misidentification Affects Biodiversity Metrics: Dealing With This Issue Using the New R Package Naturalist." *Ecological Informatics* 69: 202207.

Romero, S. M., S. L. Giudicessi, and R. G. Vitale. 2021. "Is the Fungus *Aspergillus* a Threat to Cultural Heritage?" *Journal of Cultural Heritage* 51: 107–124. https://doi.org/10.1016/j.culher.2021.08.002.

Samimi-Namin, K., Y. Benayahu, M. Ekins, et al. 2026. "A Phylogenomics-Guided Revision of the Genus *Rhytisma* Alderslade, 2000 (Octocorallia: Malacalcyonacea: Lemnaliidae), with Descriptions of Six New Species." *Invertebrate Systematics* 40: IS25068. https://doi.org/10.1071/IS25068.

Scherz, M. D., S. M. Rasolonjatova, J. Köhler, et al. 2020. "Barcode 'Fishing' for Archival DNA From Historical Type Material Overcomes Taxonomic Hurdles, Enabling the Description of a New Frog Species." *Scientific Reports* 10: 19109.

Srinivasan, M., D. Sedmak, and S. Jewell. 2002. "Effect of Fixatives and Tissue Processing on the Content and Integrity of Nucleic Acids." *American Journal of Pathology* 161, no. 6: 1961–1971. https://doi.org/10.1016/S0002-9440(10)64472-0.

Straube, N., M. Preick, G. J. P. Naylor, and M. Hofreiter. 2021. "Mitochondrial DNA Sequencing of a Wet-Collection Syntype Demonstrates the Importance of Type Material as Genetic Resource for Lantern Shark Taxonomy (Chondrichthyes: Etmopteridae)." *Royal Society Open Science* 8: 210474. https://doi.org/10.1098/rsos.210474.

Tulloch, R. L., J. Rojahn, L. E. Neaves, A. Trujillo-González, C. E. Holleley, and E. E. Hahn. 2026. "Evaluating the Molecular Potential and Interpretability of DNA in Historical Spirit Collection Media." *Molecular Ecology Resources* 26, no. 4: e70153. https://doi.org/10.1111/1755-0998.70153.

Vargas, S., O. Breedy, F. Siles, and H. M. Guzman. 2010. "How Many Kinds of Sclerite? Towards a Morphometric Classification of Gorgoniid Microskeletal Components." *Micron* 41: 158–164.

Vázquez-Ramírez, J., H. V. Narave Flores, and M. J. Cházaro Basañez. 2025. "Unexpected Consequences of Taxonomic Misidentification for the Conservation of an Endangered Species." *Conservation Science and Practice* 7: e70180. https://doi.org/10.1111/csp2.70180.

Voolstra, C. R., and M. Ziegler. 2020. "Adapting With Microbial Help: Microbiome Flexibility Facilitates Rapid Response to Environmental Change." *BioEssays* 42: 2000004. https://doi.org/10.1002/bies.202000004.

West, J. M. 1997. "Plasticity in the Sclerites of a Gorgonian Coral: Tests of Water Motion, Light Level, and Damage Cues." *Biological Bulletin* 192: 279–289. https://doi.org/10.2307/1542721.

Wood, D. E., J. Lu, and B. Langmead. 2019. "Improved Metagenomic Analysis With Kraken 2." *Genome Biology* 20: 257. https://doi.org/10.1186/s13059-019-1891-0.

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Zhang, Y., and J. J. Wiens. 2026. "Cryptic Species Are Widespread Across Vertebrates." *Proceedings of the Royal Society B* 293: 20252377. https://doi.org/10.1098/rspb.2025.2377.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Sequencing results and genomic resources generated for type specimens and other historical specimen vouchers of octocorals. See sheet 2 for definitions of all variables and acronyms. **Table S2:** Reference genomes from GenBank included in custom database used for Kraken2 analyses. **Table S3:** Figure 3-support models: omnibus tests for protocol-age models. **Table S4:** Paired genome-skimming versus target-capture comparison. **Table S4A:** Summary statistics. **Table S4B:** Specimen-level values. **Table S5:** Age-class sensitivity analyses comparing genome-skimming and target-capture in specimens <50 and ≥ 50 years old. **Table S5A:** Sample sizes by age class and sequencing protocol. **Table S5B:** Protocol × age-class interaction tests in the full dataset. **Table S5C:** Posthoc GS versus TC contrasts within age classes for outcomes with significant interactions. **Table S5D:** Ethanol-only sensitivity analysis. **Table S6:** Figure 5-support models: omnibus tests for preservation-age models. **Table S7:** Museum-adjusted models. **Table S8:** Contamination and authenticity models. **Table S9:** Generic identity of Symbiodiniaceae reads recovered from octocoral types. Total number of reads and percent of all Symbiodiniaceae reads reported for each type specimen. Symbiodiniaceae_only_%_zoox = % of Symbiodiniaceae reads that could not be identified to genus. **Table S10:** SSU rRNA sequences assembled from raw read files of octocoral types using phyloFlash. **Table S11:** phyloFlash contaminant analyses. **Table S12:** Identity of UCE loci recovered from octocoral type specimens placed in spurious 'outgroup' clade in phylogenetic analysis (Figure S3). **Figure S1:** Number of (a) PE150 reads, (b) UCE loci and (c) mitochondrial protein-coding genes recovered from octocoral type specimens as a function of the amount of DNA used for library preparation. Blue: genome-skimming; gold: target-capture. **Figure S2:** Number of UCE loci recovered from octocoral type specimens that were sequenced with both genome-skimming (blue, GS) and target-capture (gold, TC). Lines connect the same specimen. **Figure S3:** Maximum likelihood tree of Octocorallia constructed from UCE loci (50% occupancy matrix). Black: reference sequences from Quattrini et al. (2024); blue: type specimens with >100 UCE loci recovered; red: type specimens recovered in incorrect or spurious clades; purple: type in unexpected phylogenetic position confirmed by subsequent re-assessment of morphology. Types are labelled using basionym and sample ID (see Table S1). **Figure S4:** Maximum likelihood tree of Octocorallia constructed from UCE loci (75% occupancy matrix) with non-octocoral UCE loci removed from samples with spurious placement in Figure S3. Black: reference sequences from Quattrini et al. (2024); blue: type specimens with >100 UCE loci recovered; red: type specimens recovered in incorrect or spurious clades; green: type specimens recovered in correct clades after removal of non-octocoral UCEs; purple: type in unexpected phylogenetic position confirmed by subsequent reassessment of morphology. Types are labelled using basionym and sample ID (see Table S1). **Figure S5:** Maximum likelihood tree of Octocorallia constructed using mtMutS. Black: published reference sequences; blue: type specimens recovered in expected clades; red: type specimens recovered in incorrect clades; purple: type specimen recovered in unexpected phylogenetic position but confirmed by subsequent re-assessment of morphology. Types are labelled using basionym and sample ID (see Table S1). **Figure S6:** Maximum likelihood tree of Octocorallia constructed using 28S rDNA. Black: published reference sequences; blue: type specimens recovered in expected clades; red: type specimens recovered in incorrect clades; purple: type specimen recovered in unexpected phylogenetic position but confirmed by subsequent re-assessment of morphology. Types are labelled using basionym and sample ID (see Table S1).

17550998, 2026, 8, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.70236 by Cochrane Portugal, Wiley Online Library on [19/11/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License