Title: What's left in the tank? Identification of non-ascribed aquarium's coral collections with DNA barcodes as part of an integrated diagnostic approach

Authors: Luigi Colin[1]; Daniel Abed-Navandi[2]; Dalia A. Conde[3,4]; Jamie Craggs[5]; Rita da Silva[3]; Max Janse[6]; Björn Källström[7]; Alexander Pearce-Kelly[8]; Chris yesson[1]

Institutions:

1) Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK

2) Haus des Meeres - Aqua Terra Zoo, Fritz Gruenbaum Platz 1, 1060 Vienna, Austria

3) Interdisciplinary Centre on Population Dynamics (CPop), University of Southern Denmark, Biology Department, Syddansk Universitet, Campusvej 55, DD-5230 Odense M, Denmark

4) Conservation Science Alliance, Species360, 7900 International Drive, Suite 1040 Bloomington, MN, 55425 USA

5) Horniman Museum & Gardens, 100 London Road, Forest Hill, London SE23 3PQ, UK

6) Royal Burgers' Zoo, Antoon van Hooffplein 1, 6816 SH Arnhem, the Netherlands

7) Maritime Museum & Aquarium, Karl Johansgatan 1-3, SE-41459, Göteborg, Sweden

8) Zoological Society of London, Regent's Park, London NW1 4RY, UK

Communicating author: luigi.colin@ioz.ac.uk

Keywords: Coral, DNA Barcoding, Aquarium, Identification, Blast match, Species

Abstract

The unprecedented threats to coral reef ecosystems from global climate change (GCC) require an urgent response from the aquarium community, which is becoming an increasingly vital coral conservation resource. Unfortunately, many hermatypic corals in aquaria are not identified to species level, which hinders assessment of their conservation significance. Traditional methods of species identification using morphology can be challenging, especially to non-taxonomists. DNA barcoding is an option for species identification of Scleractinian corals, especially when used in concert with morphology based assessment. This study uses DNA barcodes to try to identify aquarium specimens of the diverse reef-forming family *Acroporidae* from 127 samples. We identified to species name 44% of the analysed samples and provided provisional identification for 80% of them (101/127, in the form of a list of species names with associate confidence values). We highlighted a sampling bias in public nucleotide sequences repertories (e.g.: GenBank) towards more charismatic and more studied species, even inside a well-studied genus like *Acropora*. In addition, we showed a potential "single observer" effect with over a quarter of the reference sequences used for these identifications coming from the same study. We propose the use of barcoding and query matching as an additional tool for taxonomic experts and general aquarists to increase their chances of making high confidence species-level identifications. We produce a standardised and repeatable methodology to increase the capacity of aquariums and other facilities to assess non-ascribed species.

Introduction

It is widely acknowledged that coral reef ecosystems are facing unprecedented threats from global climate change (GCC) e.g.(Pratchett et al. 2013; Hoegh-Guldberg et al. 2017; Cowburn et al. 2018; Hughes et al. 2018; Masson-Delmotte et al. 2018; Brondizio et al. 2019; Sheppard et al. 2020). Warm water reef-building corals create the most biodiverse ecosystems on the planet and have been

estimated to support 830'000 species of multi-cellular plants and animals worldwide (Bellwood and Hughes 2001; Mora et al. 2008, 2011; Fisher et al. 2015), providing a variety of habitats for fish, invertebrates and other taxa in shallow tropical seas (Bellwood and Hughes 2001). Corals in the family *Acroporidae* play a key functional ecosystem role as they are the major reef builders in the majority of warm water reefs ecosystems (Fukami et al. 2000).

Corals in aquaria

Given the severity of the species and ecosystem level climate change threats to corals (Masson-Delmotte et al. 2018), and their reliant biodiversity, *ex situ* conservation assistance is likely to play an important part in ensuring the survival of many, perhaps the majority, of coral species. Aquaria institutions therefore have a vitally important conservation remit to support coral restoration strategies allowing for both sexual and asexual coral recruitment, biobanking and a wide range of research, education and thus contribution towards *in situ* conservation activities. Early attempts in the 1980s at keeping corals in aquaria faced difficulties (Borneman 2008; Brunner 2012) it was until 1980 that *Acropora* was the the first genus to be successfully cultured by Stüber (Borneman 2008). From the 1990s onwards coral husbandry boomed, reaching a level where virtually all families of zooxanthellate corals were not only being maintained for many years, but were being propagated and traded between private and public aquaria (Borneman 2008). Specimens conserved in Aquaria thus, became a highly valuable resource for restoration strategies (Rinkevich 1995).

Identification of corals

Coral taxonomy is a traditionally difficult discipline, subject to same challenges of modern taxonomy (Godfray 2002). The genus *Acropora* is among the most diverse and geographically widespread reef building corals (Wallace 1999). Research on *Acropora* suggests that species limits

may be sometimes narrower, sometimes broader, than generally perceived (Wallace and Willis 1994). Currently the number of officially recognised *Acropora* species are between 135 (WORMS (Horton et al. 2020)) and 163 (Corals of the World (Veron et al. 2020)), reaching 186 if we include the "*taxon inquirendum"* on WORMS (Horton et al. 2020). There have been different attempts at grouping these species within the genus *Acropora* (Wallace 1999). These attempts lead to the creation of the concept of species group (aka syngameon). The concept of species groups, first defined for convenience of identification and without implying taxonomic relation, has been revised to reflect the evolution and phylogeny of the genus (Wallace 1999; Wallace et al. 2012). The use of genetic tools for species identification has provided another line of evidence for species delimitation, further adding to the blurred boundaries between species groups, species and sub-species.

Traditionally, identification of coral has relied on morphological skeletal features, with the great majority of taxa originally described following typological species concepts defined around a century ago (Best et al. 1984; Todd 2008). The relationship between morphological variation within-species and the environment has been a source of contention since the late 1800s (Todd 2008). Whether this variability is due to different underlying genotypes or plastic phenotypes is unclear. Evidence supports both scenarios (difference in phenotype e.g.: (Willis 1985; Ayre and Willis 1988; Todd et al. 2004)– Plastic phenotypes e.g.: (Foster 1979; Miller 1994; Bruno and Edmunds 1997; Muko et al. 2000; Todd et al. 2004)) and these scenarios are not necessary mutually exclusive but probable to operate simultaneously (Foster 1979; Amaral 1994; Todd 2008). This has particular relevance in the face of continuous GCC, that is creating an increasingly unpredictable wild reef environment.

Notwithstanding the taxonomic difficulties, the identification of coral species can be challenging even in optimal 'wild' conditions. However, identification requirements are far more than just for

well documented field collections, but also cover aquaria collections, that in some cased are held in conditions dramatically different from the wild. Furthermore customs officials rely in proper taxonomic identification even at the genera level for threatened taxa by the international wildlife trade (AC25 Doc. 23 CITES 2011). A status of these identification difficulties in aquaria can be demonstrated by an analysis of Species360 Zoological Information Management Software (ZIMS), adopted by more than 95 aquariums in 24 countries. Assuming correct identification ZIMS indicates c. 42.9% of corals of the orders Scleractinia, Alcyonacea, Helioporacea, Antipatharia, Corallimorpharia, Pennatulacea, and Zoanthidea in aquaria are identified to genus or higher taxonomic level (ZIMS list of species holdings – 4th June 2020). Most of the records correspond to the order Scleractinia, which is also the order with more species represented in aquariums. Corallimorpharia is, however, the order with the highest number of reported individuals in aquariums members of the Species360 network (Fig. 1). In addition to this the non-species associated genera (e.g.: *Acropora* sp.) is usually recorded once in the ZIMS database and most other aquarium inventory systems, which possibly significantly underestimates the actual number of non-identified specimens. Moreover, there are currently no confidence values ascribed to any taxonomic level a collection allocates to a specimen which means that there is likely to mean that a significant number of species ascribed specimens are actually at the lower end of the confidence spectrum and would benefit from being reassessed.

In many aquariums' corals come from three sources: fragments from other aquaria, via ornamental trade, or from confiscated and seized shipments (which aquariums are usually asked to accommodate). All three sources present identification challenges as often exact knowledge from their wild origin is absent. This specially a challenge for confiscated specimens since they have none or relatively poor source records, sample integrity, and any other additional background information that might be of assistance for the identification process. This, together with the difficulties caused by environmental phenotypic plasticity makes coral identification in aquariums a particularly challenging task. Data from the Convention on the International Trade of Endangered

Species of Flora and Fauna (CITES) shows a total of 4557 instances of coral confiscation from 1982 to 2018 worldwide, for a total of 27027 individual items and 5461 kg per year (CITES data - https://trade.cites.org 02/07/2020). Of these, less than 20% are reported at species level, meaning that for most cases we know only the genus, family or even order of the specimens confiscated. Given these challenges CITES requires ideally identification to the order level.

DNA Barcoding

An alternative or complementary diagnostic option to morphological identification is genetic analysis. DNA Barcodes have been proposed by Hebert et al. (Hebert et al. 2003) as a way to circum-vent the "limitations inherent in morphology-based identification systems". The core concept of DNA barcoding is the use of a standard sequence corresponding to a single homologous region that can be amplified with universal primers that is sufficiently variable to distinguish between species – usually represented by the mitochondrially encoded cytochrome c oxidase I (MT-CO1 or CO1 or COX1) (Hebert et al. 2003).

DNA barcoding is especially attractive as an option for species identification of Scleractinian, as this method bypasses the problems caused by phenotypic plasticity or life stage (larva, juvenile or adult). Unfortunately, the proposed 'universal' Barcoding region CO1 is at least 10-20 times slower in Cnidaria than the standard vertebrate mtDNA rate (Van Oppen et al. 1999; Shearer et al. 2002) with zero interspecific divergence for many species (Shearer and Coffroth 2008). This has led to the use of a number of different barcoding regions for corals (Shearer et al. 2002).

The diverse, reef forming Acropora group are widely grown in aquaria and have been a target for barcoding work. Potential barcodes were reviewed by Shearer et al. (Shearer et al. 2002) examining both nuclear and mitochondrial regions. Substitution rates of mitochondrial genes in *Acropora* tend to be lower relative to the nuclear counterparts mentioned by Shearer (i.e. Pax-C intron; internal transcribed spacer 1 and 2 (ITS) (Shearer et al. 2002). Although, Van Oppen et al. (van Oppen et al.

1999) and Vollmer et al. (Vollmer 2002) have highlighted the mitochondrial putative control region (mtCR) is more variable than the other mitochondrial regions. Subsequently Shearer et al. (Shearer et al. 2002) compared the genetic divergence of mitochondrial and nuclear regions in Scleractinian, highlighting how the mtCR region has a similar range of divergence to nuclear coding regions like the Pax-C intron, but is much slower than the hyper-variable ITS1-2.

This study seeks to build on these studies by investigating the utility of DNA barcoding as a method for species identification of *Acropora* collections in aquaria as part of an integrated diagnostic approach.

Material and Methods

Study Area and Collection

As part of the "idCoral" project we collected 224 samples across seven institutions (Aquarium de Paris (France), Chester's Zoo (UK), Haus des Meeres-Vienna (Austria), Horniman Museum and Gardens (UK), Royal Burgers' Zoo (Netherlands), Tierpark Hagenbeck Aquarium (Germany), ZSL London Zoo (UK)). Branch tip fragments of 2 cm were collected and stored in single ethanol (+95%) vials labelled with a unique identifier. During the sample collection the colonies were tagged and photographed with collection dates and tank locations were recorded. The photographs and tags were also recorded in the "idCoral" database (database - idcoral.org (2020)). After collection samples were stored at -20°C until extraction, with an initial change of ethanol after 1-2 days aimed at reducing degradation. We analysed 127 *Acropora* samples from the above mentioned collection.

DNA Extraction, Amplification and Sequencing

The DNA extractions were undertaken using the DNeasy PowerSoil Pro Kit from Qiagen© with a modified version of the manufacturer's protocol. An incubation step with proteinase K was added and the physical breaking of the cell structure with a chemical dissolution was removed from the protocol as it resulted in excessive shearing of the DNA (Online Resource 1). Extracted DNA was then preserved at -20°C, and the DNA concentration and quality tested with both Nanodrop assay and gel electrophoresis. The DNA was diluted to standardise the concentration to between 5 to 20 µg/ml.

Two genes, PaxC and mtCR were selected as barcodes based on an assessment of their variability and availability as reference sequences (Van Oppen et al. 2001; Shearer et al. 2002). The ITS1-2 region was excluded due to the low availability of reference sequences. PCR amplification was conducted with Q5® High- Fidelity DNA Polymerase mixed following manufacturer's indication. 3µl of DNA template and 0.5 µl of Bovine Serum Albumin (BSA)1µg were added, reaching a final volume of 25µl with RNA free water. For the PaxC gene the primer used were PaxC FP1 5'-TCC AGA GCA TTA GAG ATG CTG G-3' and PaxC RP1 5'-GGC GAT TTG AGA ACC AAA CCT GTA-3' (Van Oppen et al. 2001) with a protocol of 98°C for 30s, followed by 31 cycles at 98°C for 10 s, 62°C for 30 s and 72°C for 60 s, ending with a final phase of 72°C for 2min. For the mtCR the primer used were RNS2: 5'-CAG AGT AAG TCG TAA CAT AG-3' and GR: 5'-AAT TCC GGT GTG TGT TCT CT-3' (Suzuki et al. 2008) with a protocol of 98°C for 30s, followed by 40 cycles at 98°C for 10 s, 62°C for 15 s and 72°C for 60 s, ending with a final phase of 72°C for 5min.

Successful PCR amplifications (assessed by gel electrophoresis) were sent to Eurofins Genomics for PCR cleaning and custom Sanger sequencing on both the forward and the reverse primers using cycle sequencing technology (dideoxy chain termination / cycle sequencing) on ABI 3730XL

sequencing machines. The returned sequences were manually trimmed in order to remove poor quality sections and assembled using Geneious Prime® (2019.1.2.).

Identification

Taxonomy

Both the barcode gap and specimen identifications were initially performed at the species level, but were also performed at the level of species group based on the revision of Acropora by Wallace et al. (Wallace et al. 2012).

Barcoding Gap

References sequences for the two targeted regions were collated. FASTA files were created with the records matching the following query on GenBank: mtCR=“"Acroporidae"[Organism] AND mitochondrial "complete genome" OR "control region" OR "putative control region" OR mitochondrial control region”; PaxC=“"Acroporidae"[Organism] AND PaxC[gene] OR pax-c[gene]”. Two multiple alignments were performed, one for each gene, using the Geneious® algorithm (progressive pairwise aligner) and trimmed. For both the produced references alignments distance matrices were calculated using the function dist.dna (R package apex (Jombart et al. 2017)) and with different models (Raw ; JC69; K80; K81; F84; BH87; paralin; indel Y; indelblock). Successively the frequency distribution of intra/inter-specific distances was plotted. The quantile function (R package “stats") was used to define the 99; 50; 1% quantile confidence interval used in the blast match analysis. Note: the two alignments were constructed and used only for the specific purposes of examining the barcoding gap for the targeted region and determining potential confidence interval for the blast match analysis.

Blast match

Query based specimen identification with blast match (Benson 2000) was determined using a custom Blastn database. All records from the query “"Acroporidae"[Organism]” on GenBank were downloaded and included in the local database built with the packages rBLAST (Hahsler and Nagar 2019) and rentrez (Winter 2017). A local Blast match (Camacho et al. 2009) was ultimately performed. To filter for species match with identical scores a decision tree (Fig. 2) was designed relying on Bit score, Percentage identity, and number of mismatches. Matches shorter than 2/3 of the original sample sequence query were removed. The decision tree prioritised 100% pairwise identity matches were possible, then to the highest bit score, then lowest number of base mismatches. 100% pairwise identity matches to a species group rather than a single species name were prioritised over single species name matches with lower pairwise identity with higher bit score. We confronted the match from the two targeted regions to obtain a single most probable ID. To account for conflicting matches between the two region we relied on the 99; 50; 1% quantile distribution of the intraspecific distances obtained from the barcoding gap analysis. Using pairwise identity threshold as confidence intervals to discern which region offers the better match for the ID (R-scripts as Online Resource 3 and 4).

Results

From the 127 collected samples all were successfully extracted and, 107 mtCR and 100 PaxC had a successful DNA amplification, respectively 84% and 79% with a 76 samples overlap between the two region (Table 1). 348 reads were successfully sequenced, 84 mtCR sample and 90 PaxC. The median sequence lengths were 734 post assembly for the mtCR region and 649 for the PaxC region (Table 1).

Barcoding

Our GenBank query produced a total of 1348 mtCR sequences and 427 PaxC sequences. The mtCR references query, once cleaned for clones, produced a 579 sequences alignment of 914bp. The PaxC references query produced a 352 sequence alignment of 770 bp.

The mtCR alignment of 579 sequences resulted in a matrix of 335 241 pairwise genetic distances, grouped by 57 species names. 6.4% of these (21 471) are intraspecific distances, while 93.6% (313 770) are interspecific distances (Fig. 3A). The PaxC alignment of 352 sequences resulted in a matrix of 123 904 genetic distances, grouped by 55 species names. 3% of these (3 726) are intraspecific distances, while 97% (120 178) are interspecific distances (Fig. 3B). Between the two groups there is an overlap of 49 species names. Table 3 and 4 with the sources of the sequences used in these alignments.

Fig. 3 highlights in blue the intraspecific variation (sequences with the same name on GenBank) and in the light green the interspecific distances (sequences with different names on GenBank). In both Fig. 3A) mtCR and 3B) PaxC we see no discernable barcoding gap.

The quantile within species distribution produce the following distance threshold for our confidence interval: mtCR 0.027 (99%); 0.0068 (50%); 0 (1%) and PaxC 0.17 (99%); 0.014 (50%); 0 (1%).

In Fig. 4 and 5, for the mtCR and the PaxC respectively, we can see the above mentioned uneven distribution of the GenBank records towards specific species and species groups. With example like species group *humilis* showing 4 times the frequency of less represented species group like *loripes*.

Blast match

The local database for blast matching contained 510826 sequences. In Table 2 we see that 23 of our 76 samples (with both regions sequenced) were uniquely identified to a single species, while 13 samples matched to conflicting names. 20 of our samples had a unique species name match for the PaxC region but mtCR region matched to 2 or more names with all but 6 of these including the matching PaxC name. Very few samples matched to multiple species for both regions (4).

We have 47 % of samples matching to a single species name with both regions, 5% with the same identical match. While we have 42% of the matches to a single species name that are in conflict, 25% of these conflicts are resolvable (i.e.: one of the two region with a match better fitting than the other one, based on the defined confidence thresholds of each region). Similar results for the other sample matching for two or more in at least one of the two region. Out of the 76 samples, in 15 cases we observe the second region adding information to the first ones (Table 5 as Online Resource 2). Mainly from a mtCR multi name ID to a single PaxC name (i.e.: sample P269 has 7 mtCR identical matches to 7 different species names, and the PaxC matches at a lower rate to only one of them). In addition, the second region functions as backup option whenever one of the two falls too short (5/76 sample have too low-quality match to be trusted independently).

The complete list of our current best species name match, with species group information and detailed genetic info is provided as Online Resource (Table 5 – Online Resource).

Discussion

Barcoding gap

Using DNA barcodes to identify unknown samples is only achievable if a well- studied, well-sampled reference sequence database is available (Meyer and Paulay 2005), and even though

databases such as GenBank (which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI - http://www.ncbi.nlm.nih.gov/genbank) and BOLD (Barcode Of Life Data System - http://www.boldsystems.org) contain an ever increasing number of species-level identified sequences (i.e.: 216 531 829 total sequences on GenBank as April 2020), these databases are far from complete. The open nature of GenBank is destined to produce a bias towards species with a higher research interest (i.e.: Human sequences constitute 56% of the total sequences on GenBank (Benson 2000)). The genus focus of our study, *Acropora*, is one of the more commonly studied and widespread hard corals in the ocean (Wallace 1999; Fukami et al. 2000) making it one of the most covered coral genera in GenBank. The reality is that even for a widely studied group such as *Acropora* we can observe a disparity in the number of sequences available for each species name (Fig. 4 and 5). The best example of this is the delta between two species group: group *humilis* being the most sampled (251 mtCR and 61 PaxC references sequences from our query) and group *robusta* one of the least sampled (11 mtCR and 13 PaxC references sequences from our query). Note that both species group are formed by of the same number of species (8 species each (Wallace et al. 2012)). This disparity is intrinsic to the nature of these public repertoires, as species of higher interest will always be more studied than others, creating the necessity to acknowledge the potential bias when using these resources for identification purposes.

From our analysis we could not identify a clear barcoding gap in both our targeted regions (Fig. 3), and furthermore we can observe a total overlap of the intra-specific distances by the inter-specific distances. For both our targeted regions, even if more evident in the mtCR (Fig. 3A), we see that the inter-specific distances follow a distribution reminiscent of a more common and more expected barcoding distribution. We performed the same barcoding gap analysis at the higher taxonomic level using the species group information without obtaining a noticeable difference. With the lack of a clear Barcoding gap, and accounting for the extensive criticism present in the literature: For the use of distances in taxonomy and systematics (Ferguson 2002; Lee 2004; Moritz and Cicero 2004;

DeSalle et al. 2005; Prendini 2005; Knapp et al. 2005; Meyer and Paulay 2005; Cognato 2006; Hickerson et al. 2006; Little and Stevenson 2007; Meier et al. 2008); for the drawbacks of using GenBank data (Meier et al. 2006); for the evidence of different rates of evolution in Scleractinian (Van Oppen et al. 1999; Shearer et al. 2002; Kitahara et al. 2010). We cannot establish a simple distance base threshold to use for our identification purposes, leading us to depend more on the blast match.

Another potential issue with the use of these methods of identification is the source of the reference sequences. With databases like GenBank, with little or no review on the uploaded information, the question is the confidence that the initial identification of the now available reference sequences is indeed correct (James Harris 2003). In our specific case after the processing we performed for the mtCR reference alignment, out of the total 579 included 42% (242 sequences) come from a single doctoral thesis project, while 5.4% (33 sequences) are from unpublished papers (Table 3). For the PaxC 16% (58 sequences) are from unpublished papers, and another 39% (138 sequences) from a single paper (Richards et al. 2008) (Table 4). Highlighting clearly the need for more study on the subject to increase the number of potential sources, and alleviate the possible bias of a “single observer”.

Blast match

The use of query match with blast is far from being a complete solution, with 26 out of 76 samples without a single species or species group resolved ID (Table 5), we have proposed a theoretical decision tree that standardise the choice when there is a need of sorting through the output of a multi-species name blast match (Fig. 2). In addition we used two different regions, a mitochondrial one and a nuclear one to increase the confidence in our results (Suzuki et al. 2008). Our proposed species ID are provisional and subject to revision when more species and sequences are available, hence the value of a repeatable process for matching and re-matching.

All methods of identification that relay on DNA barcodes (e.g.: Blast matching, Barcoding Gap analysis, etc.) use genetic divergence thresholds to assign individuals to correct species. This based on two fundamental assumptions 1) monophyly of species with respect of the molecular marker used, and 2) intraspecific genetic divergence is much smaller than interspecific genetic differences (Toffoli et al. 2008). The first assumption is more dependent on the experimental choices made (i.e. sequence region selected). The second assumption is where most of the criticism of barcodes methods lies, on not comparing sister species or geographical distribution thus, underestimating intraspecific genetic variability (Moritz and Cicero 2004; Toffoli et al. 2008). Fundamentally, the issue of using thresholds lies in the fact that species naturally embodies an evolutionary process, being subject to demographic and selective processes that will act on the genetic diversity (e.g. (Avise et al. 2000; Hey 2001; Coyne and Orr 2004; Toffoli et al. 2008)). Since species are real evolutionary groups and not categories which are created as a direct function of perceived distinction (Hey 2001), the use of thresholds in "discovering" new species is overly simplistic, and in some cases even misleading (Toffoli et al. 2008). We argue against simply substituting morphology based taxonomy with DNA barcoding based taxonomy, as each system is more adapt to answer different questions (Toffoli et al. 2008). Correct assignment is thus only possible by complementing DNA barcoding with other data types, such as morphological, and ecological characters as part of an integrated diagnostic approach (Toffoli et al. 2008).

While the simple blast match is inheritably less complex than other traditional DNA barcodes identification methods (i.e.: the above mentioned barcoding gap analysis and, the potentially more direct Automatic Barcoding Gap Discovery (ABGD) (Puillandre et al. 2012)). The argument we make is that given the shared limitation and the common need for additional data in an integrated approach, a faster and simpler blast match provides a relatively similar additional amount of information for the species identification while being more easily repeatable, especially considering

the value standardized data among aquariums (da Silva et al. 2019) and in light of the ever increasing quantity of reference materials being added to GenBank and BOLD.

The reality is that barcoding as an independent tool of identification has its clear limitation (Van Oppen et al. 1999; Shearer et al. 2002; Ferguson 2002; Moritz and Cicero 2004; Lee 2004; DeSalle et al. 2005; Meyer and Paulay 2005; Prendini 2005; Knapp et al. 2005; Cognato 2006; Hickerson et al. 2006; Meier et al. 2006, 2008; Little and Stevenson 2007; Kitahara et al. 2010), and coral barcoding is no exception. With some of the concern on the use of genetic repertories in common of both the methods explored in this study, we strongly recommend an integrated diagnostic approach, combining morphological and genetic means of identification, in agreement with C. Moritz in "DNA Barcoding: Promise and Pitfalls" (Moritz and Cicero 2004). Proposing the use of barcoding and query matching as an additional tool for identification, increasing confidence of experts and both confidence and capacity of non-experts taxonomist. Producing this way reliably identified sequences that can potentially become new reference, adding to the repertories already freely available. Considering the specific needs of the aquarium community and their critical role to support corals conservation (i.e., by hosting at least 24% of the corals assessed as highly vulnerable to climate change (da Silva et al. 2019)), we emphasise the value of integrating standardised barcoding analysis, and specimen morphology photographs to optimise usage of authoritative identification guides and expert opinion.

Acknowledgments

Genetic analysis costs were partly covered by a FAITAG genetics fund grant an initiative from the EAZA (European Association of Zoos and Aquaria) and EUAC (European Union of Aquarium Curators) Fish and Aquatic Invertebrate Taxon Advisory Group (FAITAG).

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Fig. 1: Number of records from member institutions of the Species360 network at genus or higher level (i.e. order, family and genus) and at species or lower level (i.e. species and subspecies). This information is based on the "rank" field from the holdings information reported by each institution through the ZIMS software. Information downloaded from ZIMS (Species360, 2020) on 4th June 2020.

Fig. 2: Decision tree for conflict solving based on Pairwise percentage identity, Bit score and number of mismatches.

Fig. 3: Barcoding gap (raw distance matrix frequency distribution). A) mtCR max distance = 0.0877 minimum distance = 0 (Bin size=0.002) B) PaxC max distance = 0.5895 minimum distance = 0 (Bin size =0.05)

Fig. 4: Number of mtCR references sequences with the same species name by species group

Fig. 5:Number of PaxC references sequences with the same species name by species group

Bibliography

Amaral FD (1994) Morphological variation in the reef coral *Montastrea cavernosa* in Brazil. Coral Reefs 13:113–117

Avise JC, Baker RJ, Porter CA, Patton JC, Van Den Bussche RA, Ballard JW, Chemoff B, James AC (2000) Phylogeography: the history and formation of species. Harvard university press,

Ayre DJ, Willis BL (1988) Population structure in the coral *Pavona cactus*: clonal genotypes show little phenotypic plasticity. Mar Biol 99:495–505

Bellwood DR, Hughes TP (2001) Regional-scale assembly rules and biodiversity of coral reefs. Science (80-) 292:1532–1534

Benson DA (2000) GenBank. Nucleic Acids Res 28:15–18

Best, Borel M, G. J. Boekschoten, Oosterbaan A (1984) Species concept and ecomorph variation in living and fossil Scleractinia. Palaeontogr Am 54 70–79

Borneman E (2008) Introduction to the husbandry of corals in aquariums : A review. Adv coral Husb public aquariums, Leewis, RJ Janse, M 2:3–14

Brondizio ES, Settele J, Dìaz S, Others (2019) IPBES. 2019 Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Bonn, Ger

Brunner B (2012) The Ocean at Home: an illustrated history of the aquarium. Reaktion Books,

Bruno JF, Edmunds PJ (1997) Clonal variation for phenotypic plasticity in the coral Madracis mirabilis. Ecology 78:2177–2190

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: Architecture and applications. BMC Bioinformatics 10:1–9

Cognato AI (2006) Standard Percent DNA Sequence Difference for Insects Does Not Predict Species Boundaries. J Econ Entomol 99:1037–1045

Cowburn B, Samoilys MA, Obura D (2018) The current status of coral reefs and their vulnerability to climate change and multiple human stresses in the Comoros Archipelago, Western Indian Ocean. Mar Pollut Bull 133:956–969

Coyne JA, Orr HA (2004) Speciation Sinauer Associates. Sunderland, MA 276:281

DeSalle R, Egan MG, Siddall M (2005) The unholy trinity: Taxonomy, species delimitation and DNA barcoding. Philos Trans R Soc B Biol Sci 360:1905–1916

Ferguson JWH (2002) On the use of genetic divergence for identifying species. Biol J Linn Soc 75:509–516

Fisher R, O'Leary RA, Low-Choy S, Mengersen K, Knowlton N, Brainard RE, Caley MJ (2015) Species richness on coral reefs and the pursuit of convergent global estimates. Curr Biol 25:500–505

Foster AB (1979) Phenotypic plasticity in the reef corals *Montastraea annularis* (Ellis & Solander) and *Siderastrea siderea* (Ellis & Solander). J Exp Mar Bio Ecol 39:25–54

Fukami H, Omori M, Hatta M (2000) Phylogenetic Relationships in the Coral Family *Acroporidae*, Reassessed by Inference from Mitochondrial Genes. Zoolog Sci 17:689–696

Godfray HCJ (2002) Challenges for taxonomy. Nature 417:17–19

455 Hahsler M, Nagar A (2019) rBLAST:Interface to the Basic Local Alignment Search Tool (BLAST).

456 Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA
457 barcodes. Proc R Soc B Biol Sci 270:313–321

458 Hey J (2001) Genes, categories, and species: The evolutionary and cognitive cause of the species
459 problem. Oxford University Press,

460 Hickerson MJ, Meyer CP, Moritz C (2006) DNA barcoding will often fail to discover new animal
461 species over broad parameter space. Syst Biol 55:729–739

462 Hoegh-Guldberg O, Poloczanska ES, Skirving W, Dove S (2017) Coral Reef Ecosystems under
463 Climate Change and Ocean Acidification. Front Mar Sci 4:158

464 Horton T, Kroh A, Ahyong S, Bailly N, Boyko CB, Brandão SN, Gofas S, Hooper JNA, Hernandez
465 F, Holovachov O, Mees J, Molodtsova TN, Paulay G, Decock W, Dekeyzer S, Lanssens T,
466 Vandepitte L, Vanhoorne B, Adlard R, Adriaens P, Agatha S, Ahn KJ, Akkari N, Alvarez B,
467 Anderberg A, Anderson G, Angel M V, Antic D, Arango C, Artois T, Atkinson S, Baldwin
468 BG, Bank R, Barber A, Barbosa JP, Bartsch I, Bellan-Santini D, Bergh N, Bernot J, Berta A,
469 Bezerra TN, Bieler R, Blanco S, Blasco-Costa I, Blazewicz M, Bock P, de León M, Böttger-
470 Schnack R, Bouchet P, Boury-Esnault N, Boxshall G, Bray R, Bruce NL, Cairns S, Calvo
471 Casas J, Carballo JL, Cárdenas P, Carstens E, Chan BK, Chan TY, Cheng L, Churchill M,
472 Coleman CO, Collins AG, Collins GE, Corbari L, Cordeiro R, Cornils A, Coste M, Costello
473 MJ, Crandall KA, Cremonte F, Cribb T, Cutmore S, Dahdouh-Guebas F, Daly M, Daneliya M,
474 Dauvin JC, Davie P, De Broyer C, De Grave S, de Mazancourt V, de Voogd NJ, Decker P,
475 Decraemer W, Defaye D, d'Hondt JL, Dippenaar S, Dohrmann M, Dolan J, Domning D,
476 Downey R, Ector L, Eisendle-Flöckner U, Eitel M, Encarnação SC d., Enghoff H, Epler J,
477 Ewers-Saucedo C, Faber M, Feist S, Figueroa D, Finn J, Fišer C, Fordyce E, Foster W, Frank
478 JH, Fransen C, Freire S, Furuya H, Galea H, Gao T, Garcia-Alvarez O, Garcia-Jacas N, Garic
479 R, Garnett S, Gasca R, Gaviria-Melo S, Gerken S, Gibson D, Gibson R, Gil J, Gittenberger A,
480 Glasby C, Glover A, Gómez-Noguera SE, González-Solís D, Gordon D, Gostel M, Grabowski
481 M, Gravili C, Guerra-García JM, Guidetti R, Guiry MD, Gutierrez D, Hadfield KA, Hajdu E,
482 Hallermann J, Hayward BW, Heiden G, Hendrycks E, Herbert D, Herrera Bachiller A, Ho J s.,
483 Hodda M, Høeg J, Hoeksema B, Houart R, Hughes L, Hyžný M, Iniesta LFM, Iseto T,
484 Ivanenko S, Iwataki M, Janssen R, Jarms G, Jaume D, Jazdzewski K, Jersabek CD, Jóźwiak P,
485 Kabat A, Kantor Y, Karanovic I, Karthick B, Katinas L, Kim YH, King R, Kirk PM, Klautau
486 M, Kociolek JP, Köhler F, Kolb J, Kotov A, Kremenetskaia A, Kristensen RM, Kulikovskiy
487 M, Kullander S, Lambert G, Lazarus D, Le Coze F, LeCroy S, Leduc D, Lefkowitz EJ,
488 Lemaitre R, Lichter-Marck IH, Liu Y, Loeuille B, Lörz AN, Lowry J, Ludwig T, Lundholm N,
489 Macpherson E, Madin L, Mah C, Mamo B, Mamos T, Manconi R, Mapstone G, Marek PE,
490 Marshall B, Marshall DJ, Martin P, Mast R, McFadden C, McInnes SJ, Meidla T, Meland K,
491 Merrin KL, Messing C, Miljutin D, Mills C, Moestrup Ø, Mokievsky V, Monniot F, Mooi R,
492 Morandini AC, da Rocha R, Morrow C, Mortelmans J, Mortimer J, Musco L, Nesom G,
493 Neubauer TA, Neubert E, Neuhaus B, Ng P, Nguyen AD, Nielsen C, Nishikawa T, Norenburg
494 J, O'Hara T, Opresko D, Osawa M, Osigus HJ, Ota Y, Páll-Gergely B, Panero JL, Pasini E,
495 Patterson D, Paxton H, Pelser P, Peña-Santiago R, Perrier V, Perrin W, Petrescu I, Picton B,
496 Pilger JF, Pisera AB, Polhemus D, Poore GC, Potapova M, Pugh P, Read G, Reich M, Reimer
497 JD, Reip H, Reuscher M, Reynolds JW, Richling I, Rimet F, Ríos P, Rius M, Rogers DC,
498 Roque N, Rosenberg G, Rützler K, Sabbe K, Saiz-Salinas J, Sala S, Santagata S, Santos S, Sar
499 E, Satoh A, Saucède T, Schatz H, Schierwater B, Schilling E, Schmidt-Rhaesa A, Schneider S,
500 Schönberg C, Schuchert P, Senna AR, Serejo C, Shaik S, Shamsi S, Sharma J, Shear WA,
501 Shenkar N, Shinn A, Short M, Sicinski J, Sierwald P, Simmons E, Sinniger F, Sivell D, Sket
502 B, Smit H, Smit N, Smol N, Souza-Filho JF, Spelda J, Sterrer W, Stienen E, Stoev P, Stöhr S,
503 Strand M, Suárez-Morales E, Summers M, Suppan L, Susanna A, Suttle C, Swalla BJ, Taiti S,

Tanaka M, Tandberg AH, Tang D, Tasker M, Taylor J, Taylor J, Tchesunov A, ten Hove H, ter Poorten JJ, Thomas JD, Thuesen E V, Thurston M, Thuy B, Timi JT, Timm T, Todaro A, Turon X, Tyler S, Uetz P, Uribe-Palomino J, Urtubey E, Utevsky S, Vacelet J, Vachard D, Vader W, Väinölä R, Valls Domedel G, de Vijver B, van der Meij SE, van Haaren T, van Soest RW, Vanreusel A, Venekey V, Vinarski M, Vonk R, Vos C, Walker-Smith G, Walter TC, Watling L, Wayland M, Wesener T, Wetzel CE, Whipps C, White K, Wieneke U, Williams DM, Williams G, Wilson R, Witkowski A, Witkowski J, Wyatt N, Wylezich C, Xu K, Zanol J, Zeidler W, Zhao Z (2020) World Register of Marine Species (WoRMS).

Hughes TP, Connolly SR, Eakin CM, Dietzel A, McWilliam MJ, Hoey AS, Pears RJ, Liu G, Skirving WJ, Heron SF, Pratchett MS, Baird AH, Stella JS, Hoogenboom MO, Torda G, Kerry JT (2018) Global warming transforms coral reef assemblages. Nature 556:492–496

James Harris D (2003) Can you bank on GenBank? Trends Ecol Evol 18:317–319

Jombart T, Kamvar ZN, Schliep K, Archer E, Harris R (2017) apex: Phylogenetic Methods for Multiple Gene Data.

Kitahara M V., Cairns SD, Stolarski J, Blair D, Miller DJ (2010) A comprehensive phylogenetic analysis of the *scleractinia (Cnidaria, Anthozoa)* based on mitochondrial CO1 sequence data. PLoS One 5:

Knapp S, Lughadha EN, Paton A, Mallet J, Isaac NJB, MacE GM, James Harris D, Froufe E (2005) Taxonomic inflation, species concepts and global species lists (multiple letters). Trends Ecol Evol 20:7–8

Lee MSY (2004) The molecularisation of taxonomy. Invertebr Syst 18:1–6

Little DP, Stevenson DW (2007) A comparison of algorithms for the identification of specimens using DNA barcodes: Examples from gymnosperms. Cladistics 23:1–21

Masson-Delmotte V, Zhai P, Pörtner H-O, Roberts D, Skea J, Shukla PR, Pirani A, Moufouma-Okia W, Péan C, Pidcock R, Connors S, Matthews JBR, Chen Y, Zhou X, Gomis MI, Lonnoy E, Maycock T, Tignor M, Waterfield T (2018) IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, an. Sustain Dev Efforts to Eradicate Poverty https//www ipcc.ch/sr15/ 1:

Meier R, Shiyang K, Vaidya G, Ng PKL (2006) DNA barcoding and taxonomy in diptera: A tale of high intraspecific variability and low identification success. Syst Biol 55:715–728

Meier R, Zhang G, Ali F (2008) The use of mean instead of smallest interspecific distances exaggerates the size of the “barcoding gap” and leads to misidentification. Syst Biol 57:809–813

Meyer CP, Paulay G (2005) DNA barcoding: Error rates based on comprehensive sampling. PLoS Biol 3:e422

Miller KJ (1994) Morphological variation in the coral genus *Platygyra* - Environmental influences and taxonomic implications. Mar Ecol Prog Ser 110:19–28

Mora C, Tittensor DP, Adl S, Simpson AGB, Worm B (2011) How many species are there on earth and in the ocean? PLoS Biol 9:1–8

Mora C, Tittensor DP, Myers RA (2008) The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. Proc R Soc B Biol Sci 275:149–155

Moritz C, Cicero C (2004) DNA Barcoding: Promise and Pitfalls. PLoS Biol 2:e354

548 Muko S, Kawasaki K, Sakai K, Takasu F, Shigesada N (2000) Morphological plasticity in the coral
549 *Porites sillimaniani* and its adaptive significance. Bull Mar Sci 66:225–239

550 van Oppen MJH, Hislop NR, Hagerman PJ, Miller DJ (1999) Gene Content and Organization in a
551 Segment of the Mitochondrial Genome of the *Anthozoan Subclass Zoantharia*. Mol Biol Evol
552 16:1812–1815

553 Van Oppen MJHH, McDonald BJ, Willis B, Miller DJ (2001) The evolutionary history of the coral
554 genus *acropora (Scleractinia, Cnidaria)* based on a mitochondrial and a nuclear marker:
555 Reticulation, incomplete lineage sorting, or morphological convergence? Mol Biol Evol
556 18:1315–1329

557 Van Oppen MJHH, Willis BL, Miller DJ (1999) Atypically low rate of cytochrome b evolution in
558 the *scleractinian* coral genus *Acropora*. Proc R Soc London Ser B Biol Sci 266:179–183

559 Pratchett MS, Pisapia C, Sheppard CRC (2013) Background mortality rates for recovering
560 populations of *Acropora cytherea* in the Chagos Archipelago, central Indian Ocean. Mar
561 Environ Res 86:29–34

562 Prendini L (2005) Comment on "Identifying spiders through DNA barcodes." Can J Zool 83:498–
563 504

564 Puillandre N, Lambert A, Brouillet S, Achaz G (2012) ABGD, Automatic Barcode Gap Discovery
565 for primary species delimitation. Mol Ecol 21:1864–1877

566 Richards ZT, van Oppen MJH, Wallace CC, Willis BL, Miller DJ (2008) Some rare Indo-Pacific
567 coral species are probable hybrids. PLoS One 3:1–8

568 Rinkevich B (1995) Restoration Strategies for Coral Reefs Damaged by Recreational Activities:
569 The Use of Sexual and Asexual Recruits. Restor Ecol 3:241–251

570 Shearer TL, Coffroth MA (2008) DNA BARCODING: Barcoding corals: limited by interspecific
571 divergence, not intraspecific variation. Mol Ecol Resour 8:247–255

572 Shearer TL, van Oppen MJH, Romano SL, Wörheide G (2002) Slow mitochondrial DNA sequence
573 evolution in the *Anthozoa (Cnidaria)*. Mol Ecol 11:2475–2487

574 Sheppard C, Sheppard A, Fenner D (2020) Coral mass mortalities in the Chagos Archipelago over
575 40 years: Regional species and assemblage extinctions and indications of positive feedbacks.
576 Mar Pollut Bull 154:111075

577 da Silva R, Pearce-Kelly P, Zimmerman B, Knott M, Foden W, Conde DA (2019) Assessing the
578 conservation potential of fish and corals in aquariums globally. J Nat Conserv 48:1–11

579 Suzuki G, Hayashibara T, Shirayama Y, Fukami H (2008) Evidence of species-specific habitat
580 selectivity of *Acropora* corals based on identification of new recruits by two molecular
581 markers . Mar Ecol Prog Ser 355:149–159

582 Todd PA (2008) Morphological plasticity in *scleractinian* corals. Biol Rev 83:315–337

583 Todd PA, Sidle RC, Lewin-Koh NJI (2004) An aquarium experiment for identifying the physical
584 factors inducing morphological change in two massive *scleractinian* corals. J Exp Mar Bio
585 Ecol 299:97–113

586 Toffoli D, Hrbek T, de Araújo MLG, de Almeida MP, Charvet-Almeida P, Farias IP (2008) A test
587 of the utility of DNA barcoding in the radiation of the freshwater stingray genus *Potamotrygon*
588 (*Potamotrygonidae, Myliobatiformes*). Genet Mol Biol 31:324–336

589 Veron JEN, Stafford-Smith MG, Turak E, L.M. D (2020) Corals of the World.
590 http://www.coralsoftheworld.com

591 Vollmer S V. (2002) Hybridization and the Evolution of Reef Coral Diversity. Science (80-)
592 296:2023–2025

593 Wallace CC (1999) Staghorn corals of the world: a revision of the coral genus *Acropora*
594 (*Scleractinia; Astrocoeniina; Acroporidae*) worldwide, with emphasis on morphology,
595 phylogeny and biogeography. CSIRO publishing.

596 Wallace CC, Done BJ, Muir PR (2012) Revision and catalogue of worldwide staghorn corals
597 *acropora* and *isopora* (scleractinia: *Acroporidae*) in the museum of tropical Queensland. Mem
598 Queensl Museum 57:1–257

599 Wallace CC, Willis BL (1994) Systematics of the coral genus *Acropora*: implications of new
600 biological findings for species concepts. Annu Rev Ecol Syst 25:237–262

601 Willis BL (1985) Phenotypic plasticity versus phenotypic stability in the reef corals *Turbinaria*
602 *mesenterina* and *Pavona cactus*. Proc Fifth Int Coral Reef Symp Vol. 4:

603 Winter DJ (2017) rentrez: An R package for the NCBI eUtils API. R J 9:520–526

604 (2020) idCoral.org. https://idcoral.org

605

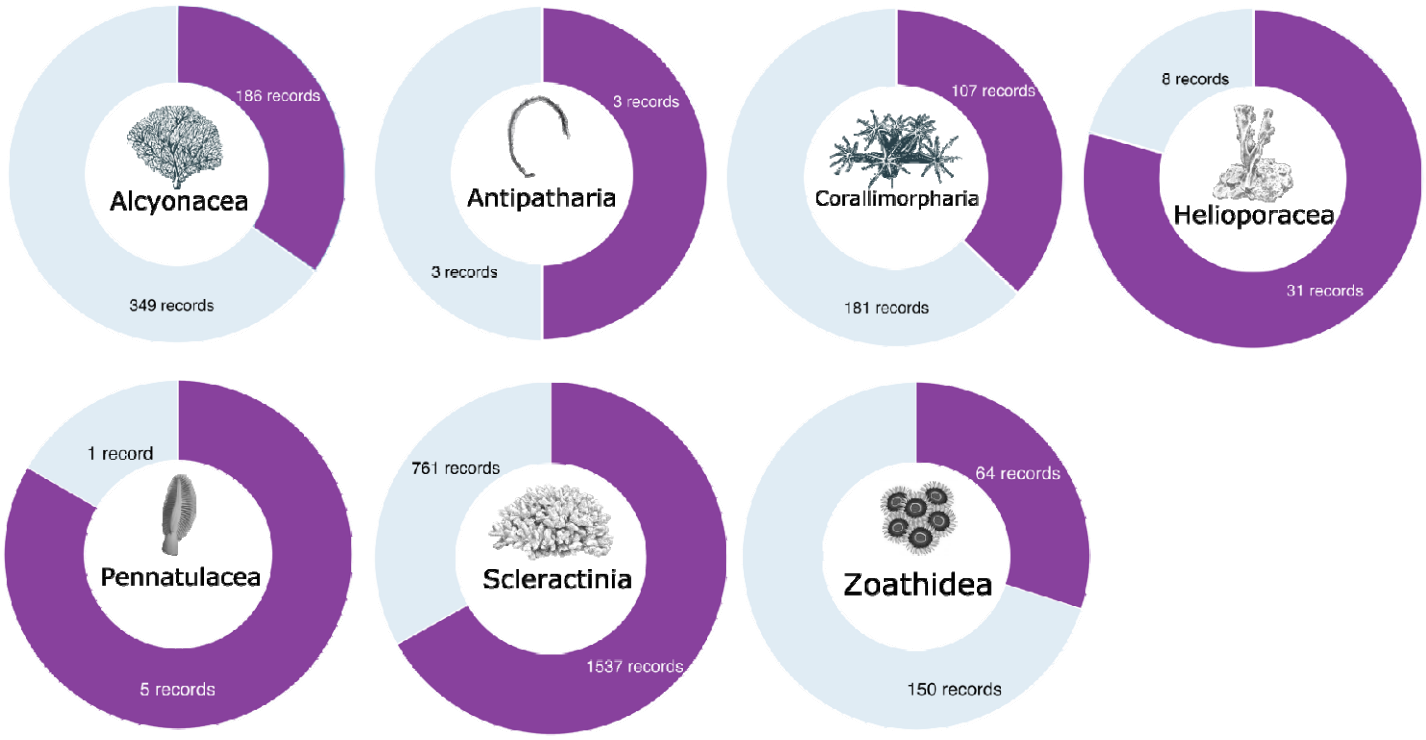

Figure 1: Number of records from member institutions of the Species360 network at genus or higher level in light blue(i.e. order, family and genus) and at species or lower level in purple (i.e. species and subspecies). This information is based on the rank field from the holdings information reported by each institution through the ZIMS software. Information downloaded from ZIMS (Species360, 2020) on 4th June 2020.

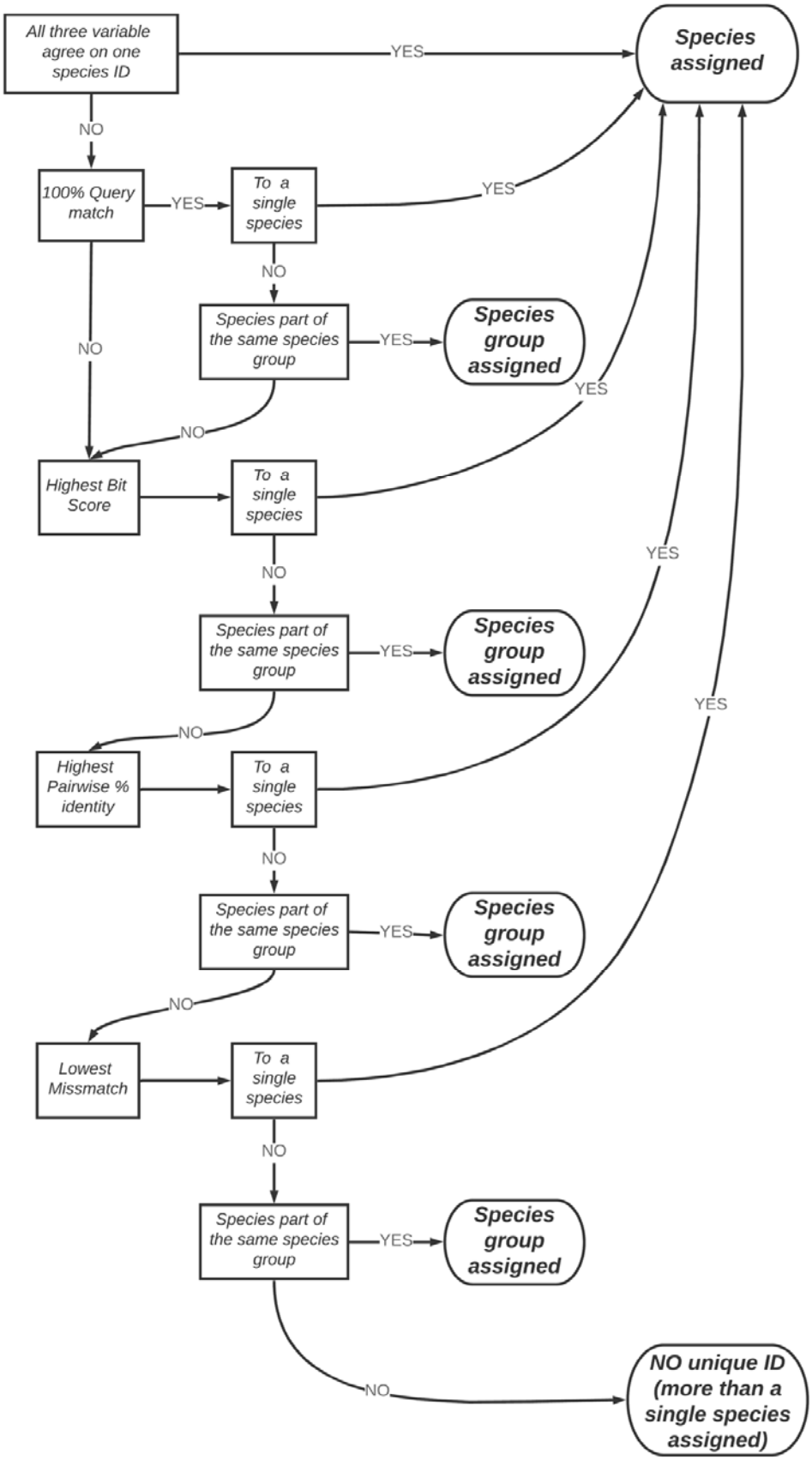

Figure 2: Decision tree for conflict solving based on Pairwise percentage identity, Bit score and number of mismatches.

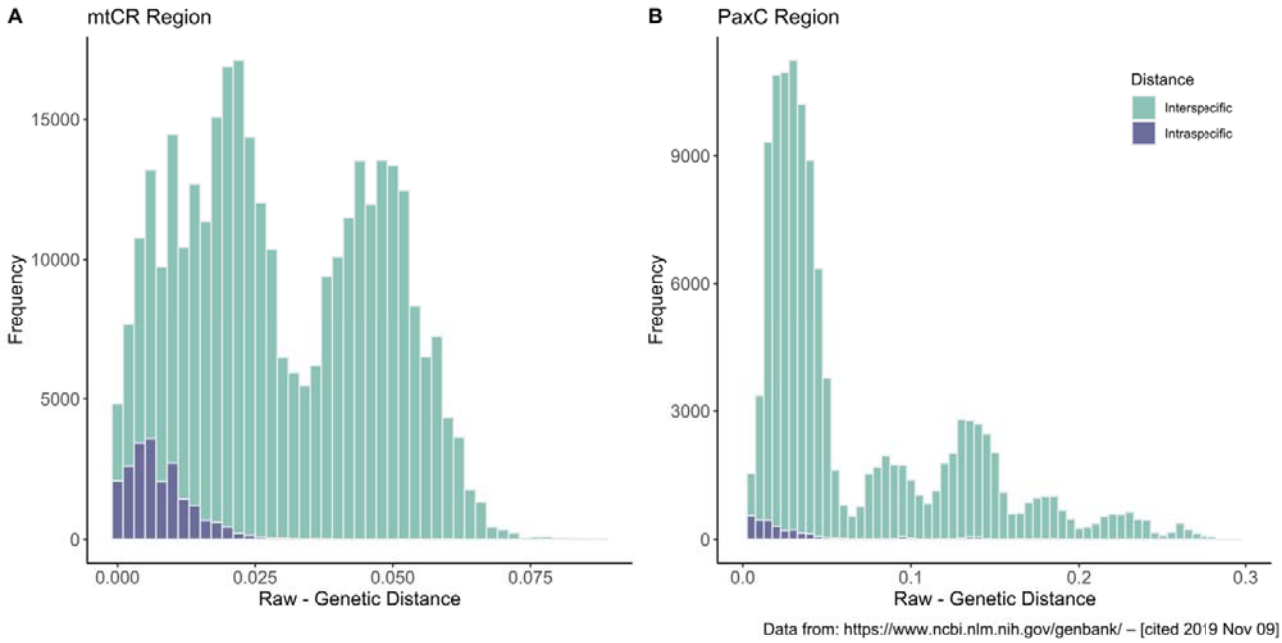

Figure 3: Barcoding gap (raw distance matrix frequency distribution). A) mtCR max distance = 0.0877 minimum distance = 0 (Bin size=0.002) B) PaxC max distance = 0.5895 minimum distance = 0 (Bin size =0.05)

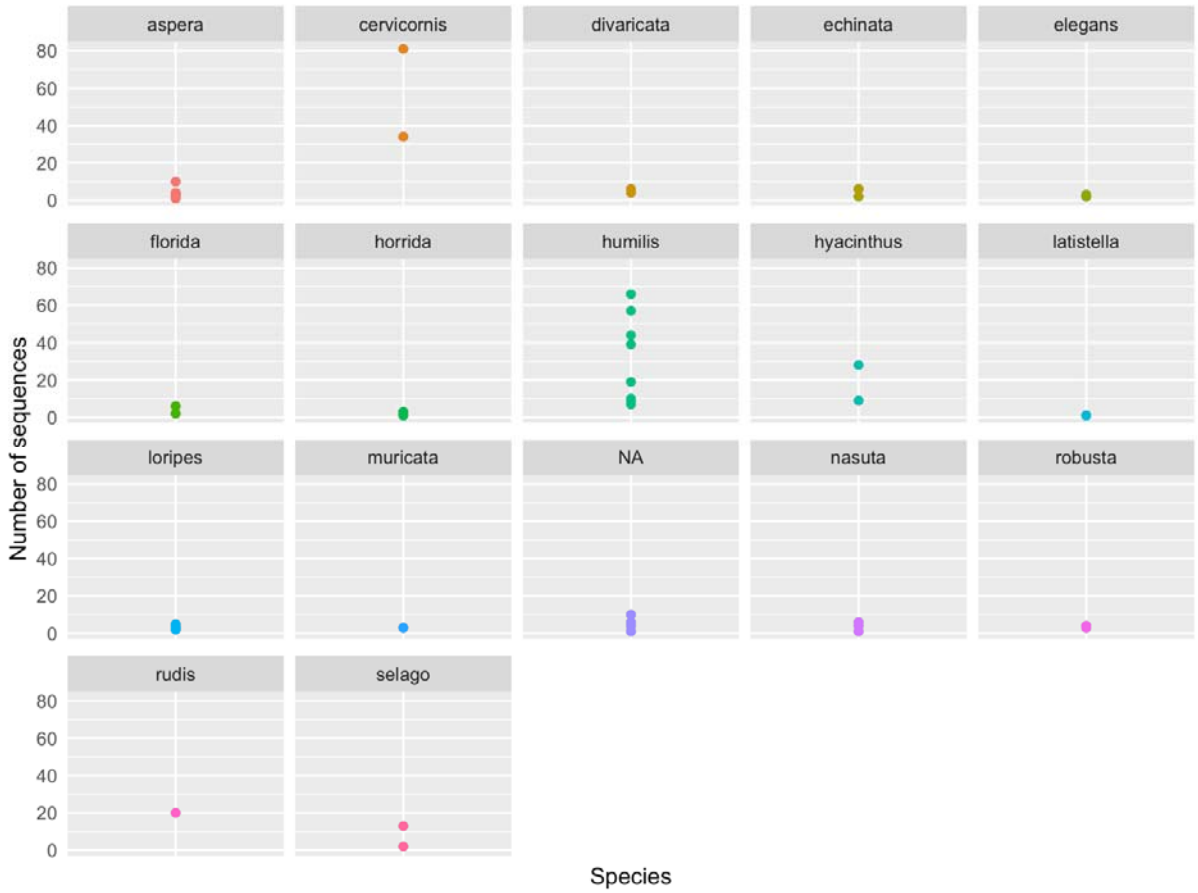

Figure 4: Number of mtCR references sequences with the same species name by species group

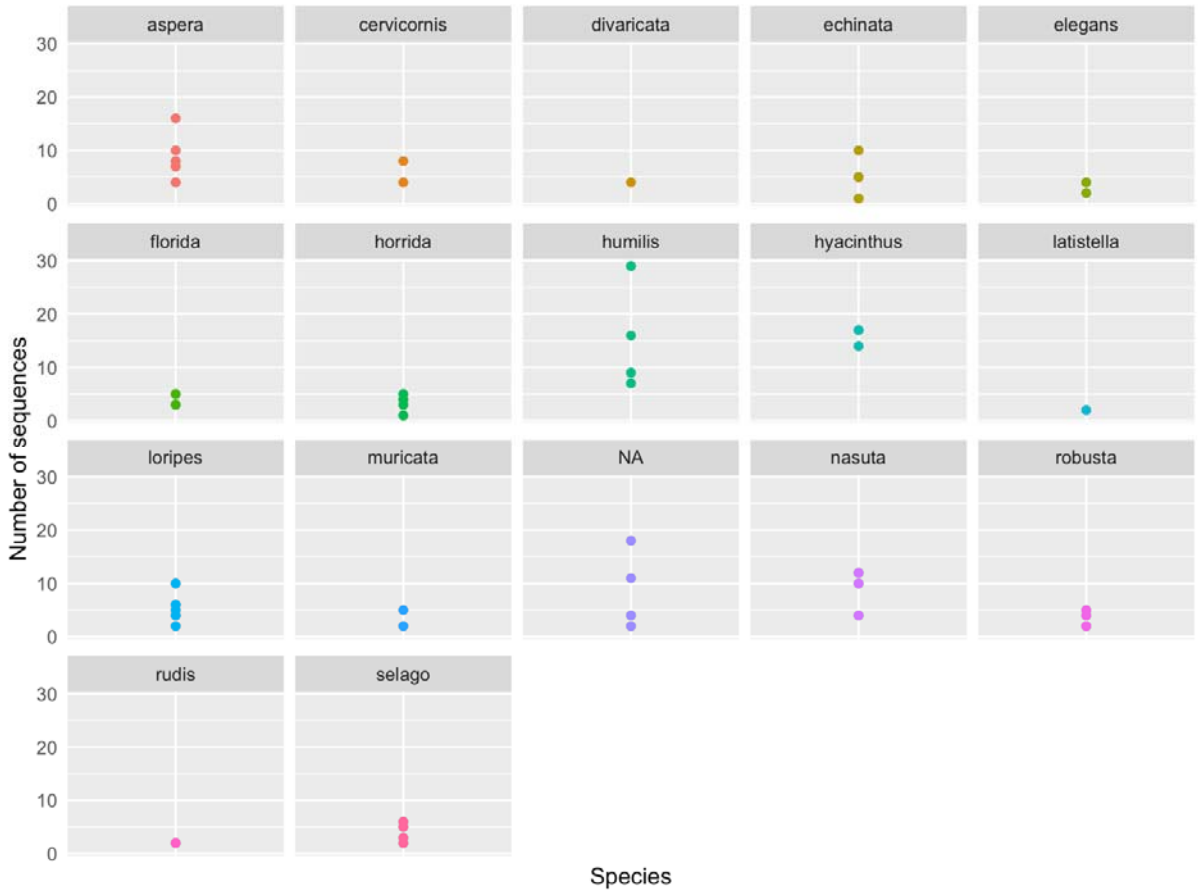

Figure 5: Number of PaxC references sequences with the same species name by species group

Table 1: Success rate across genes

N samples	N samples overlapping	Gene	N Sequences	Minimum sequence length	Maximum sequence length	Median sequence length
127	76	*mtCR*	84	563	904	734
		PaxC	90	429	869	649

Table 2: Number of sample matching to a single or more (2+) species name on GenBank. Above the diagonal line match that agree between the two region. Below the diagonal line matches that disagree between the two region

		mtCR	
		1	2+
PaxC	1	23 / 13	20 / 6
	2+	6 / 4	1 / 3

Table 3:Sources of mtCR references sequences by Paper title

Journal	Authors	Title	N	%
Doctroal Thesis (2003) James Cook University, Townsville, Queensland, Australia	Wolstenholme JK	Species boundaries in scleractinian corals: a case study of the Acropora humilis species group	242	41.8
PLoS ONE 3 (9), E3240 (2008)	Z. T. Richards, M. J. H. van Oppen, C. C. Wallace, B. L. Willis, and D. J. Miller	Some rare Indo-Pacific coral species are probable hybrids	79	13.6
Conserv. Genet. 18 (4) 825–835 (2017)	M. D. Waterhouse, C. Blair, K. W. Larsen, and M. A. Russello	Genetic Variation and Fine-Scale Population Structure in the Threatened Acropora palmata and Acropora cervicornis around Puerto Rico	48	8.3
Science 296 (5575), 2023-2025 (2002)	S. V. Vollmer	Hybridization and the evolution of reef coral diversity	42	7.3
Mol. Biol. Evol. 18 (7), 1315-1329 (2001)	M. J. H. van Oppen, B. J. McDonald, B. Willis, and D. J. Miller	The evolutionary history of the coral genus Acropora (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence?	41	7.1
PLoS One 5 (1) e8652 (2010)	E. M. Hemond and S. V Vollmer	Genetic Diversity and Connectivity in the Threatened Staghorn Coral (Acropora cervicornis) in Florida	32	5.5
Mol. Ecol. 11 (8), 1339-1349 (2002)	L. M. Márquez, M. J. H. Van Oppen, B. L. Willis, A. Reyes, and D. J. Miller	The highly cross-fertile coral species, Acropora hyacinthus and Acropora cytherea, constitute statistically distinguishable lineages	21	3.6
Zool. Sci. 29 (2), 134-140 (2012)	G. Suzuki and H. Fukami	Evidence of genetic and reproductive isolation between two morphs of subtropical-dominant coral Acropora solitaryensis in the non-reef region of Japan	13	2.2
Mol. Ecol. 24 (19), 5006-5019 (2015)	N. L. Rosser	Asynchronous spawning in sympatric populations of a hard coral reveals cryptic species and ancient genetic lineages	8	1.4
PeerJ 7, e6429 (2019)	H. Fukami, K. Iwao, N. H. Kumagai, M. Morita, and N. Isomura	Maternal inheritance of F1 hybrid morphology and colony shape in the coral genus Acropora	6	1.0
Syst. Biodivers. 8 (2), 281–288 (2010)	Z. T. Richards, C. C. Wallace, and D. J. Miller	Archetypal 'elkhorn' coral discovered in the Pacific Ocean	6	1.0
Mol. Phylogenet. Evol. 63 (2), 527-531 (2012)	Y. Nakajima, A. Nishikawa, A. Iguchi, and K. Sakai	The population genetic approach delineates the species boundary of reproductively isolated corymbose acroporid corals	4	0.7
Mar. Ecol. Prog. Ser. 355, 149–159 (2008)	G. Suzuki, T. Hayashibara, Y. Shirayama, and H. Fukami	Evidence of species-specific habitat selectivity of Acropora corals based on identification of new recruits by two molecular markers	3	0.5
J. Mol. Evol. 55 (1), 1-13 (2002)	M. J. H. Van Oppen, J. Catmull, B. J. McDonald, N. R. Hislop, P. J. Hagerman, and D. J. Miller	The mitochondrial genome of Acropora tenuis (Cnidaria; Scleractinia) contains a large group I intron and a candidate control region	1	0.2
Mitochondrial DNA, 1-2 (2015) In press	Y. Zhang, X. Yu, Z. Zhou, and B. Huang	The complete mitochondrial genome of Acropora aculeus (cnidaria, scleractinia, acroporidae)	1	0.2
Proc. R. Soc. Lond., B, Biol. Sci. 266 (1415), 179-183 (1999)	M. J. H. Van Oppen, B. L. Willis, and D. J. Miller	Atypically low rate of cytochrome b evolution in the scleractinian coral genus Acropora	1	0.2
Unpublished	Santacruz-Castro,A. and Dai,C.F.	Latitudinal cline in the reproductive traits and local adaptation of the Acropora hyacinthus species complex	20	3.5
Unpublished	Chan,C.L. and Chen,C.A.	Multiplex next generation sequencing of scleractinian mitochondrial genomes	11	1.9

Table 4:Sources of PaxC references sequences by Paper title

Journal	Authors	Title	N	%
PLoS ONE 3 (9), E3240 (2008)	Z. T. Richards, M. J. H. van Oppen, C. C. Wallace, B. L. Willis, and D. J. Miller	Some rare Indo-Pacific coral species are probable hybrids	138	39.2
Mol. Biol. Evol. 18 (7), 1315-1329 (2001)	M. J. H. van Oppen, B. J. McDonald, B. Willis, and D. J. Miller	The evolutionary history of the coral genus Acropora (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence?	87	24.7
Mol. Ecol. 11 (8), 1339-1349 (2002)	L. M. Márquez, M. J. H. Van Oppen, B. L. Willis, A. Reyes, and D. J. Miller	The highly cross-fertile coral species, Acropora hyacinthus and Acropora cytherea, constitute statistically distinguishable lineages	34	9.7
Mol. Ecol. 24 (19), 5006-5019 (2015)	N. L. Rosser	Asynchronous spawning in sympatric populations of a hard coral reveals cryptic species and ancient genetic lineages	16	4.5
Syst. Biodivers. 8 (2), 281–288 (2010)	Z. T. Richards, C. C. Wallace, and D. J. Miller	Archetypal 'elkhorn' coral discovered in the Pacific Ocean	10	2.8
Evol. Ecol. (2011) In press	S. R. Palumbi, S. Vollmer, S. Romano, T. Oliver, and J. Ladner	The role of genes in understanding the evolutionary ecology of reef building corals	3	0.9
Galaxea, J. Coral Reef Stud. 10 (2), 91 (2008)	M. Hatta and K. Matsushima	Presumed natural hybrids between Acropora donei and Acropora tenuis	3	0.9
Mol. Ecol. 9 (9), 1363-1373 (2000)	M. J. H. Van Oppen, B. L. Willis, H. W. J. A. Van Vugt, and D. J. Miller	Examination of species boundaries in the Acropora cervicornis group (Scleractinia, cnidaria) using nuclear DNA sequence analyses	3	0.9
Unpublished	Wei,N.V., Tang,P.-C. and Chen,A.C.	Hybridization does not occur among common Acropora species in a marginal coral assemblage	29	8.2
Unpublished	Ohki,S., Kowalski,R.K., Kitanobou,S. and Morita,M.	Different timing of the spawning is related to reproductive isolation and gamete species recognition in the broadcast spawning coral Acropora	27	7.7
Unpublished	van Oppen,M.J.H. and Miller,D.J.	A new single-locus nuclear DNA marker for evolutionary studies in the scleractinian coral genus Acropora	2	0.6