Resolving Species: Urobatis halleri, U. concentricus, and U. maculatus as Subspecies

Scott Heffernan

University of Colorado Boulder

Follow this and additional works at: https://scholar.colorado.edu/honr_theses

Recommended Citation
https://scholar.colorado.edu/honr_theses/292

This Thesis is brought to you for free and open access by Honors Program at CU Scholar. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Resolving Relationships: *Urobatis halleri*, *U. concentricus*, and *U. maculatus* as Subspecies

By
Scott Heffernan
Evolutionary Biology and Ecology, University of Colorado at Boulder

Defense Date: October 31, 2012

Thesis Advisor:
Dr. Andrew Martin, Evolutionary Biology and Ecology

Defense Committee:
Dr. Andrew Martin, Evolutionary Biology and Ecology
Dr. Barbara Demmig-Adams, Evolutionary Biology and Ecology
Dr. Nicholas Schneider, Astrophysical, Planetary, and Atmospheric Sciences
Abstract

Hybridization is the interbreeding of separate species to create a novel species (hybrid). It is important to the study of evolution because it complicates the biological species concept proposed by Ernst Mayr (1963), which is widely adopted in biology for defining species. This study investigates possible hybridization between three stingrays of the genus *Urobatis* (Myliobatiformes: Urotrygonidae). Two separate loci were chosen for investigation, a nuclear region and the mitochondrial gene NADH2. Inability to resolve three separate species within the mitochondrial phylogeny indicate that gene flow has occurred between *Urobatis maculatus*, *Urobatis concentricus*, and *Urobatis halleri*. Additionally, the lack of divergence within the nuclear gene indicates that these three species are very closely related, and may even be a single species. Further investigation is recommended with a larger sample base and additional genes.

Introduction

There are many definitions for what constitutes a species, though the most widely adopted is the biological species concept, proposed by Ernst Mayr in 1963. Under this concept, members of a species can “actually and potentially interbreed” (Mayr 1963), whereas members of different species cannot. While this concept is useful when comparing members of distantly related species, it breaks down when comparing members of closely related species (for example horses and donkeys), especially when these species have overlapping species boundaries.

Species boundaries are the geographic and genetic ranges of a species. They are defined by intrinsic barriers to gene flow, including 1) mating features (reproductive timing, mate choice, fertilization), 2) ecological factors leading to mating probability, and 3) developmental pathways which could lead to sterile offspring (Knowlton 2000). Species boundaries are often “fuzzy,” meaning there is not a clear delineation between where one species might end and another begins. For example, *Ensania eschscholtzii*, a plethodontid salamander in California, forms a geographically horseshoe-shaped range around the Central Valley. Each neighboring population can interbreed with the next, but the two populations at the terminal ends of the horseshoe cannot interbreed with each other. Under the biological species concept, these two populations would count as separate species; but it is clear that both populations belong to the same species complex (Wake 1997). In this case, the fuzzy regions are along the north-south axis of the geographic range, where the populations are in contact with one another and exchange genetic information.

Another case of fuzzy species boundaries exists in what are known as hybrid zones. Hybrid zones are the areas where the boundaries of two separate species meet and the corresponding species interbreed, producing hybrids. Although more distantly related species will not interbreed due to various factors (and therefore do not have hybrid zones), more closely related species that can interbreed will produce hybrid offspring in a hybrid zone. Depending on how much genetic divergence has occurred between the two parent species, these hybrid offspring will be sterile (more divergence) or fertile (less divergence). It is the fertile hybrid
offspring that are particularly interesting because they require us to reassess the classification of the two parent species according to the biological species concept.

Hybrid zones have been studied quite extensively in recent years (Ferris et al. 1983, Madeleine et al. 2002, Knowlton 2000, Mallet et al. 2007), and as more data are collected the familiar concept of species as set groups is receding and the idea of species as dynamic populations in constant flux is becoming more apparent. Although many of the hybrid species discovered occur in groups with external fertilization, hybridization events in elasmobranchs, which have internal fertilization, are recently being discovered (Morgan et al. 2011). Internal fertilization necessitates mate choice, which provides an additional reproductive isolation mechanism. The discovery of hybrid elasmobranchs can therefore shed light on the mating behavior of these fishes. This study investigates hybridization within a genus of stingray (Myliobatiformes; Urotrygonidae; Urobatis), which have overlapping geographic boundaries along the Pacific coast of Central America and California.

*Urobatis concentricus* is restricted to Mexico from Islas Cedros and off Rocos Chester (27°53′N, 115°04′W), central Baja California, through the Gulf of California and to Bahia Huatulco, Oaxaca, southern Mexico (IUCNredlist.org) (Figure 1). *Urobatis maculatus* is restricted to the Mexican Pacific, ranging from Lagunas Ojo de Liebre-Guerrero Negro in central Baja California (de la Cruz Agüero et al. 1996), through the Gulf of California (Figure 2). *Urobatis halleri* ranges from Humboldt Bay in northern California to Panama in Central America, but appears to be most common between southern California and Baja California (Figure 3).
Figure 1: Range distribution for *Urobatis concentricus*, from IUCNredlist.org.
Figure 2: Range distribution for *Urobatis maculatus*, from IUCNredlist.org.
All three species of the pacific *Urobatis* complex inhabit inshore, benthic habitats near bays or inlets. They are often found in eelgrass, which is utilized as camouflage, or on sandy bottoms near reefs. Of the three species, only *U. halleri* has been studied in much depth (Babel 1966), likely due to its larger range compared to *U. concentricus* and *U. maculatus*. The habitat, which is easily accessible, combined with the large, overlapping range makes *Urobatis* an ideal study genus for hybridization and species boundary penetration in elasmobranchs.

The study site chosen was in La Paz, Baja California Sur, Mexico. This site was ideal because of its location near the center of the range overlap between the three species, where hybrids would most likely be found.

Two regions of the DNA were chosen for investigation. Although many authors use cytochrome oxidase I (COI), this study focused on the mitochondrial gene NADH2 because it is more rapidly evolving (Moore et al. 2011), making it better suited for resolving recent between-species and within-species evolutionary patterns. Mitochondrial genes are also useful because they are inherited maternally, which means that they can be informative of mating patterns between species. This is discussed further below.

The second region was a microsatellite region from the nuclear gene known as Uro170. Microsatellites are useful for hybridization studies because they undergo rapid evolution due to lack of selection pressure, so more short-term patterns can
be resolved. Unlike mitochondrial genes, nuclear genes are inherited from both paternal and maternal lineages, which means it cannot yield information on mating preference.

While morphological data can be used if the species in question differ phenotypically, genetic data describes genotypic characters, which can differ greatly even when morphology is the same between two species (Babik et al. 2005, Graham et al. 1998). Phylogenetic analysis is a useful way of using morphological or genetic data to construct the most likely evolutionary tree. These trees can then be used to infer evolutionary patterns, and even hybridization (Meng and Kubatko 2009). For species that do not interbreed (no gene flow between species), phylogenetic trees will be composed of branches of separate species. However, if the species are interbreeding, genetic divergence decreases, and some taxa may group more closely with other species because they are hybrids.

Two hypotheses were formulated for this study:

$H_0$: There is no gene flow between species, and phylogenetic analysis will reveal three distinct species clusters (Figure 4).

$H_A$: There is gene flow among the species, and phylogenetic analysis will not reveal three distinct species clusters.

Figure 4: Phylogeny of *Urobatis* spp. assuming no hybridization. Simplified from Naylor et al. 2012, who found 3 species clusters.
Methods

Samples were obtained from SCRIPPS Institute and CICIMAR databases as well as from Dr. Andrew Martin’s collections, and collected in the field. Field sampling was done along Playa de Balandra in La Paz, Mexico (24°21’N 110°17’W). Specimens were spotted via snorkeling and captured in-situ with a pole-net (diameter = 43 cm, length = 2 m) on the reef. This typically necessitated a pursuit in order to capture the stingray, which could last less than five minute to upwards of half an hour. Specimens were then carried to the beach by keeping the net on the surface of the water. This was to prevent asphyxiation of the stingray during transport while keeping the stingray from escaping the net. Once on the beach, a fin clip and photo-identification was procured (See Appendix III for example). Specimens were then released back into the water. Fin-clip scars were used as identification to prevent recapture. Fin clips were stored in tubes of 95% ethanol solution and stored in a cool, dry backpack to prevent solar or temperature fouling.

DNA was extracted from each sample using Qiagen Spin-Column Protocol (Appendix I). 10 primers were chosen (Appendix II) for amplification. If a primer yielded sufficient results in gel electrophoresis, it was sent for sequencing at Functional Biosciences, Inc. Sufficiency of results was determined by strength of signals from gel electrophoresis (primers with weak signals or too few signals were repeated through PCR, using a fresh batch of DNA samples, or abandoned). Of the samples that were sequenced, only 2 had clear signals (Figure 5), which means that only two of the 10 primers were successful in amplifying a sequence.

Once the sequences were returned, contigs were formed using Sequencher. These contigs were then manually scanned and any errors corrected using the strength of the signal. This was necessary because the signal strength is typically weak at the ends of a sequence (Figure 5 b and c). However, because the two primers do not overlap evenly, where one signal is weak, the other is typically strong. Once all of the contigs were corrected, they were aligned to a reference sequence obtained from the NCBI website. The regions that were extraneous (i.e. bases were labeled as ‘N’ because the signal was too weak, or regions that had poor signals) were removed, typically cutting the sequence from ~1000 base pairs (with ~300 N bases) to ~660 base pairs of good quality (no N bases). The program PAUP was used to create a phylogenetic tree with these sequences using the Neighbor Joining method. One cladogram was produced for each sequence.

The consistency index explains how much a character has changed compared to how much possible change could have occurred. A lower consistency index means there are more actual changes compared to the number of changes that could have occurred, implying a high mutation rate.
Results

Mitochondrial gene (UroR1+F1):
The tree length was 127, meaning a total of 127 base-pair changes were necessary to explain the tree (Figure 6). The consistency index (CI) was 0.9528, and the homoplasy index was 0.0472. The retention index was 0.9794. The rescaled consistency index was 0.9331, meaning there were slightly more changes that happened compared to the estimated possible changes. This means that the mutation rate is relatively low within the taxa described.

There is a clear grouping of *U. halleri* and *U. concentricus*, with *U. maculatus* forming a sister clade with *U. concentricus*. Additionally, there are *U. maculatus* that group within the *U. halleri* and *U. concentricus* clades.

Nuclear locus (Uro170):

There was no difference found among the taxa in this study. One change occurred between the taxa collected for this study and the *U. halleri* sequence on NCBI, which was collected in California (Planck et al. 2010).

Figure 5: Sequencing results viewed in Sequencher. (a) shows a clear signal, while (b) shows a weak signal. (c) shows an unclear signal, which cannot be used.

(a)

(b)

(c)
Figure 6: Cladogram for the mitochondrial gene, NADH2. *U. jamaicensis* falls as the outgroup, and a distinct *U. halleri* cluster forms a sister clade with the *U. concentricus* and *U. maculatus* complex.
Discussion

NADH-2

The mitochondrial gene is inherited maternally, making this phylogenetic reconstruction informative of maternal genes within the population. The outgroup is *U. jamaicensis*, a sister species found in the Caribbean. The Caribbean species would be expected to be more distantly related to the three Pacific taxa, and this is what is shown in the cladogram (Figure 6). The fact that *U. concentricus* and *U. halleri* form distinct groupings implies that they are distinct mating populations (they do not mate with each other). The same can be said for the three *U. maculatus* that form a sister group with *U. concentricus*, however there are many *U. maculatus* that group within the *U. concentricus* branch and one that groups within the *U. halleri* branch. The implication is that maternal genes from *U. halleri* and *U. concentricus* are being transmitted to *U. maculatus*. The most likely cause of this transmission is the interbreeding between males of *U. maculatus* and females of both *U. concentricus* and *U. halleri*. This clearly shows that the three species are able to interbreed, meaning they are not separate species under the biological species concept.

170 Microsatellite (nuclear gene)

The lack of a difference of even 1 base pair between the taxa indicates that they are very closely related. An estimated mutation rate of $1 \times 10^{-8}$ is for nuclear genes because the mutation rate in elasmobranchs is considerably slower than in other taxa (A. Martin, pers. comm.). Our base pair length at the nuclear site is 402, so the maximum time for divergence is:

\[
\frac{1}{402} \times 10^{-8} = 248,756 \text{ years}
\]

This is a very short timescale for evolutionary processes, and it is highly unlikely for species to diverge within this time. Even for closely related species or geographically separated members of a single species, there is at least one base-pair difference in nuclear genes such as microsatellites (Costello et al. 2003).

Biological implications

The mitochondrial data show that *U. maculatus* is able to breed with both *U. concentricus* and *U. halleri*. This is evidenced in the results by the appearance of *U. maculatus* within the *U. halleri* cluster and within the *U. concentricus* cluster. Combined with the identical nuclear genes, the conclusion that there is only one species with three subspecies or types is most likely. Mate preference may preclude the mating of *U. halleri* with *U. concentricus*, as indicated by the mitochondrial gene (it is also possible that no *U. halleri/U. concentricus* hybrids were caught). The low diversity in nuclear data could also be due to a low effective population size (Hauser et al. 2002). If this is the case, there could have been a bottleneck effect within the last 25k years in the Pacific *Urobatis* complex. However, further research is needed before this can be asserted. *Urobatis* is not listed as endangered (listed as data deficient) and is not of major importance to fisheries (IUCNredlist.org), and so more study into this genus has the potential to be more informative.
Another implication is that the species status of these three subspecies needs to be revised. I propose using the species name *pacificus* due to the Pacific distribution of the taxa, and using the names *Urobatis pacificus halleri*, *Urobatis pacificus concentricus*, and *Urobatis pacificus maculatus* for the relevant subspecies names. It should also be noted that the subspecies differ (judging from the data) by coloration patterns, and genetic analysis of subspecies cannot differentiate them.

The third implication, which has an effect on the study of evolution, is that this species may be of use as a model organism for speciation by sexual preference. Since *U. concentricus* and *U. halleri* do not seem to breed with each other, they may be in the first stages of speciation (at the splitting point in an evolutionary tree). The *Urobatis* species complex would be a good model organism to because of its accessible habitat and large potential for novel information.

Potential problems

The most important problem that could have occurred in this study is a misidentification of stingrays. Because the three study species look very similar, they can easily be mistaken for one another. Except for the 19 specimens that were personally caught and identified, it is uncertain whether or not correct identifications were made. It is possible that the *U. maculatus* outliers were misidentified, and that there is no gene flow between the three species. However for this study it is assumed that the specimens were correctly identified.

The second problem is an issue with the primer used for the nuclear sequence. When the sequence was returned, it was missing the nucleotide repeats definitive of a microsatellite. The sequence was used anyway because it was missing from all of the taxa, so it should still be informative. Unfortunately, not all of the taxa came back with useable sequence data, such that the number of individuals in the nuclear tree is less than the number of individuals in the mitochondrial gene tree. The likely cause could be an issue with the primer, and if this study were done again (with a larger budget) it is recommended to develop several primers for each locus and using the best one.

Difficulties and Constraints

As mentioned briefly in the methods, data capture was done with a two-man team, snorkeling and manually catching each stingray, with only 3 full days of sampling, due to lack of reliable transportation to the study site. It was initially planned to utilize the local fishermen for more samples, but the stingray fishing season had ended right after arrival to La Paz.

The biggest constraint was monetary. Although I had a UROP grant, it lasted only enough to get me to Mexico and back, and then there was a little left for sequencing. With a larger budget, I would have stayed in the field longer, utilized more resources (boats, excursions to different sites, personnel, etc.), and been able to sequence more data (sequencing was costly because we had to send them to a private company, Functional Biosciences Inc.).

Time was another constraint. With a larger budget, I would certainly have stayed in the field for longer and sampled from more sites. By the end of my stay I had the contacts and experience that would have been beneficial for a longer (2+
months) field study. I did have enough time to complete the lab work, although with more primers I would have needed more time as well.

Further Research

With more genetic regions and a larger sample size, the interrelationships between these three species will become more apparent. Therefore I recommend further inquiry with a larger database. Additional tests for incomplete lineage sorting would also be helpful in determining the evolutionary history of these organisms. Experimental designs to test the interspecies mating capabilities of these three species would answer the question of mating preference and fertility of interspecies copulations. Laboratory observations of interspecies offspring might also give clues as to why there is no mixing of mitochondrial DNA between U. halleri and U. concentricus.

References


Appendix
I. Qiagen Spin-Column Protocol
   Perform all centrifugation steps at room temperature
   If necessary, redissolve any precipitates in Buffers ATL and AL
   Ensure that ethanol has been added to Buffers AW1 and AW2
   Preheat a thermomixer, shaking water bath, or rocking platform for heating at 56˚C
   If using frozen tissue, equilibrate the sample to room temperature

Procedure
1. Cut tissue (up to 25 mg; op to 10mg spleen) into small pieces, and place in 1.5 ml microcentrifuge tube. Add 180µl Buffer ATL.
2. Add 20 µl proteinase K. Mix by vortexing, and incubate at 56°C until completely lysed. Vortex occasionally during incubation, or place in a thermomixer in a shaking water bath, or on a rocking platform.
3. Vortex for 15 s. Add 200µl Buffer AL to the sample. Mix thoroughly by vortexing. Then add 200µl ethanol (96%-100%). Mix again thoroughly.
4. Pipet the mixture into a DNeasy Mini spin column in a 2 ml collection tube. Centrifuge at ≥6000 x g (8000 rpm) for 1 min. Discard flow-through and collection tube.
5. Place the spin column in a new 2 ml collection tube. Add 500 µl Buffer AW1. Centrifuge for 1 min at ≥6000 x g. Discard flow-through and collection tube.
6. Place the spin column in a new 2 ml collection tube. Add 500 µl Buffer AW2. Centrifuge for 3 min at 20000 x g (14000 rpm). Discard flow-through and collection tube.
7. Transfer the spin column to a new 1.5 ml or 2 ml microcentrifuge tube, and add 200µl [150 µl was used during this experiment] Buffer AE for elution. Incubate for 1 min at room temperature. Centrifuge for 1 min at ≥6000 x g. Recommended: Repeat this step for maximum yield [Repeated 2 times].

II. Primers used for PCR
1) Uha176F 5’-ACCTTGAATTTCCTTTGGG
   Uha176R 5’-TAAACCAGTTCCAGTGAGG

2) Uha170F 5’-TTCTGCAACAAACCAAAACATG
   Uha170R 5’-CAGAACCAACAGTCCAG

3) Uha115F 5’-GCTAAGCCAGAGTGGAGG
   Uha115R 5’-AAATGTGTAATTCCCTTTGG

4) Uha111F 5’-TGCTGCATAATGCTTGTTTG
   Uha111R 5’-TAAAGCTTTGCAGAGGGATT

5) Uha61F 5’-TAGGCTCCATTTGCTTGACT
   Uha61R1 5’-CTTCCCCGCACACAAATAAT
   Uha61R2 5’-TCAATGCCCAAGATCTCCTC
6) Uha36F 5’-CATGGGACTTGACAACCCTAA
   Uha36R 5’-AAGTGAGGCTTCCATTGGTG

7) Uha20F 5’-GCTCCCTGACCTTTTGCATCT
   Uha20R 5’-CCTTTCCAACCAAAAGAACA

8) Ilem 5’-AAGGACCACCTTTGATAGGT
   Asn 5’-AACGCTTAGCTGTTAATTAA

9) UroF1 5’-GAGCAACAGAAGCAACCACA
   UroR1 5’-AAATTGTGGCCTGGGATGAAG

10) HCO – Primer sequence lost
    LCO – Primer sequence lost

III. Results from PAUP-constructed trees

**NADH2**
Tree length: 127
Consistency Index (CI): 0.9528
Homoplasy index (HI): 0.0472
CI excluding uninformative characters: 0.8750
HI excluding uninformative characters: 0.1250
Retention index: 0.9794
Rescaled consistency index: 0.9331

IV. Example identification picture of CO2, the second *Urobatis concentricus*
specimen caught. A 30 cm ruler was placed next to the stingray in each photo for
size comparison.