Investigating the genetic diversity, population structure and demography of Baja and Pacific *Odontodactylus scyllarus*

Ethan Holleman

Introduction

Odontodactylus scyllarus, or the Mantis shrimp are a unique group of carnivorous stomatopods that are believed to have diverged from members of the class *Malacostraca* around 340 million years ago [1]. Mantis shrimp are notable for both their unique mechanically driven raptorial smashing or spearing claws, which are powerful enough to induce cavitation in the surrounding water [2, 3] and acutely sensitive visual system which is believed to be the most complex ever discovered in nature [4, 5]. Due to their extraordinary physiology, the biology of Mantis shrimp species has been the subject interdisciplinary study by researchers in the fields of optics, material science and marine biology. However, comparatively few studies have focused on the genetic structure of Mantis shrimp populations and have focused on Asian Mantis shrimp species in the Yellow and East China Seas [6]. Mantis shrimp are known to play an important role in marine ecosystems by acting as efficient predators of other crustaceans and oxygenating ocean sediments [7]. Understanding the genetic history of these crustaceans is therefore important for accessing the overall health of important marine ecosystems, such as coral reefs which *O. scyllarus* are known to inhabit. Further, Mantis shrimp are a significant seafood resource that while part of Asain cusin is not as often consumed in United States could prodide insignts into future fishery mangement as seafood indusry continues to expand. Furture conservation efforts of an iconic and physiologically outlier in nature.

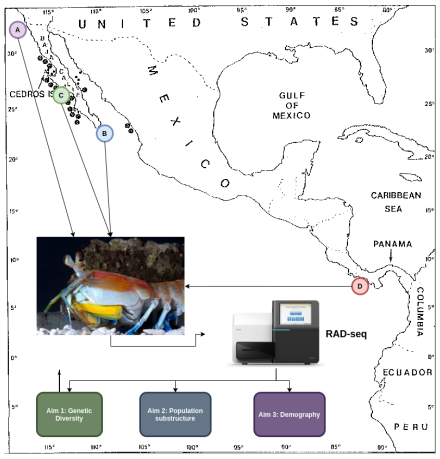

Figure 1: Proposed *Hemisquilla californiensis* collection locations shown as colored dots. Image adapted from *Basch et al, 93*

Therefore, our main goal is to quantify genetic diversity of *O. scyllarus* across their Atlantic range in order to access gene flow between populations and infer effects of recent antropogenic changes to the marine ecosystem by evaluating *O. scyllarus* demographic and selective events. Towards this effort we will pursue three specific aims.

Aim 1: Characterize genetic variation.

Genetic data will be gathered used restriction site associated DNA sequencing (RAD-seq) [8] which will allow the low-cost genotyping of a large number of individuals. RAD-seq libraries will then be used to calculate population level statistics useful towards accessing gene flow between populations namely pairwise inbreeding F-statistics. To quantify genetic differentiation

This needs to be more about D and that kind of thing

Aim 2: Quantify *O. scyllarus* population sub-structure.

Hypothesis: *Odontodactylus scyllarus* will show significant gene flow between sampling locations. Previous studies examining genetic variation of *Oratosquilla oratoria*, a Mantis shrimp species closely related to the California Mantis shrimp, in the Yellow and East China seas found a surprising degree of gene flow between populations as determined by pairwise F_{st} evaluations [6]. However, it is currently unknown if Baja and Pacific *O. scyllarus* exhibit similarly high levels of gene flow. To evaluate this question, we will collect *O. scyllarus* individuals from four main locations; three along the the Western coast of Baja California, and one on the Western coast of Panama which is believed to be *O. scyllarus* most Southern range. This will allow for comparisons to be drawn about the degree of gene flow between Northern and Southern populations. Using RAD-seq data collected in aim 1, we will calculate population level statistics

useful towards accessing gene flow between populations namely inbreeding F-statistics. First we will determine reductions in heterozygosity due to genetic differentiation between the sampled populations by calculating the pairwise F_{st} values. Ultimately F_{st} is evaluated against the expected heterozygosity of the sub-populations if they were merged, had equal size and were randomly mating. Additionally, F_{st} assumes Hardy-Weinberg equilibrium of the evaluated populations in order to access the heterozygosity from allele frequencies calculated from genotypic data. We will evaluate error in our estimations through permutation testing, which repeatedly randomly assigns individuals to a sub-population and evaluates F_{st} to ultimately create a probability distribution which can then be used to evaluate the significance of our actual results.

We will further evaluate sub-population structure using a Bayesian approach to sub-population assignment. This method assigns individuals to a sub-population given their genotype and the allele frequencies of each sub-population under the assumption of a lack of linkage disequilibrium (all loci are treated as independent). Since the origin population of each individual is known, predictive capacity can be assessed to determine this kin

Last thing here could be admixture analysis which need to watch these videos about

MAybe change this to admixture analysis

Aim 2: Evaluate sub-population structure within sampling locations

Mantis shrimp species are known to be mostly solitary creatures, rarely spending significant amounts of time outside of their sea floor burrows [10]. This small range is expected to give rise to significant population sub-structure within the four sampling locations.

- Fixation Index: reduction in genetic diversity of sub populations due to differentiation among sub populations.
  - Access the reduction of differentiation to compare specific sub-populations using pairwise comparison.
  - How to define different sub populations? Blobs of areas / habitats
  - High pairwise Fst to gauge the distinctness of populations. Do something where defining subpopulations at greater and greater distances to determine how quickly Fst will drop off at different sampling sites.
- Another entry in the list

Expected heterozygotes Ht (total population was randomly mating) Collect genotype data from a single locus A/T SNP and collect adults Determine allele frequencies of locus interested in This is where would look at mean Fst (fixation index which is mean reduction in Hs (expected heterozygotes due to genetic differentiation among sub-populations)). Fis (inbreeding coefficient reduction in Hi due to non-random mating within a sub-populations) Bayesian sub-population assignment possibly admixture analysis here as well

Aim 3: Determine demographic history between Northern and Southern California Mantis shrimp populations.

Hypothesis: *O. scyllarus* will display a higher coalescent effective population size than instantaneous effective population size due to recent ocean acidification.

Same magnitude of genetic drift as the actual population (Wright Fisher population)

Evaluate the effect of selection on the observed site frequency spectrum gleaned from RAD-seq data collected as a part of aim 1 by evluating the allele frequency selection of each suitable locus compared to the allele frequency spectrum of all other suitable loci across the genome.

Would expect this to be demographic event though since fast maybe too extreme? This article suggests that

For colensent need Pi and possibly mutation rate.

Instantous assumes constant size (wright fisher population) instanous effectice size is the same as the coalsenset size (same amount of drift and genetic diversity Pi is colasent effective size) need a mutation rate as well to estimate this.

Expected Pi is average pairwise nucleotide differences this is also called theta for some reason the larger a population is the larger we expect pi to be. Need to pick a specific locus and determine the number of nucleotide differences Pi is also tajima's theta which is an estimate of population size given a mutation rate.

Ocean acidification would not effect mantis shrimp because did not really affect their weapons https://www.nature.com/

- How to tell demography vs selection?
  - Wright-Fisher model predicts something. Neutral expectation. Constant size and no selection.
  - Selection
    - Selective sweep
    - negative selection
    - balancing selection
    - Any of these is occurring at specific loci that is under selection compared to a population (demographic level) change that should be observed genome wide because not specific to specific locus.
  - Historical demography of a population
    - Step 1 determine allele frequency spectrum of a population by sampling individuals and determine what that is
    - Neutral site freq spectrum has predictable decay but when looking at actual population and determine allele freq spectrum which might differ greatly from neutral expectation
    - Step 2 is build some demographic model
      - Pop size on y and time on x axis
      - Shows how population size has changed over time
      - s is the strength of the decline (or more generally the change in population size)
    - Step 3 Determine the parameters that we are going to vary
      - Might want to determine the
      - Strength of decline
      - Timing of decline
      - Rate of expansion
      - Some combination
      - These are the things that we are trying to determine.
    - Step 4: Simulate data with different model parameter values.
      - Get a simulated SFS sSFS
      - DO many many simulations where you change the variables of interest and compare simulated SFS to observed SFS to try and find the parameter valeus that produce a simulated SFS to the population we are interested in.
      - This is ABC approximate Bayesian computation
      - Requires defining a prior distribution that will define the parameter space
      - Simulate with the same amount of data that we collected and get gene copies sequences (simulated) which allows for calculating the simulated allele frequency spectrum

5038141913

References

[1] Cara Van Der Wal, Shane T. Ahyong, Simon Y.W. Ho, and Nathan Lo. The evolutionary history of stomatopoda (crustacea: Malacostraca) inferred from molecular data. *PeerJ*, 5:e3844, September 2017.

[2] S. N. Patek, W. L. Korff, and R. L. Caldwell. Deadly strike mechanism of a mantis shrimp. *Nature*, 428(6985):819–820, April 2004.

[3] S. N. Patek and R. L. Caldwell. Extreme impact and cavitation forces of a biological hammer: strike forces of the peacock mantis shrimp odontodactylus scyllarus. *Journal of Experimental Biology*, 208(19):3655–3664, October 2005.

[4] Thomas W. Cronin, Michael J. Bok, N. Justin Marshall, and Roy L. Caldwell. Filtering and polychromatic vision in mantis shrimps: themes in visible and ultraviolet vision. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1636):20130032, February 2014.

[5] Susan Milius. Life: Mantis shrimp flub color vision test: Results suggest alternate way of detecting wavelengths. *Science News*, 182(6):11–11, September 2012.

[6] Mei Yang and Xinzheng Li. Population genetic structure of the mantis shrimp oratosquilla oratoria (crustacea: Squillidae) in the yellow sea and east china sea. *Journal of Oceanology and Limnology*, 36(3):905–912, May 2018.

[7] P.J. Antony, S. Dhanya, P.S. Lyla, B.M. Kurup, and S. Ajmal Khan. Ecological role of stomatopods (mantis shrimps) and potential impacts of trawling in a marine ecosystem of the southeast coast of india. *Ecological Modelling*, 221(21):2604–2614, October 2010.

[8] RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, 10(2):108–108, February 2011.

[9] Xiaojun Zhang, Jianbo Yuan, Yamin Sun, Shihao Li, Yi Gao, Yang Yu, Chengzhang Liu, Quanchao Wang, Xinjia Lv, Xiaoxi Zhang, Ka Yan Ma, Xiaobo Wang, Wenchao Lin, Long Wang, Xueli Zhu, Chengsong Zhang, Jiquan Zhang, Songjun Jin, Kuijie Yu, Jie Kong, Peng Xu, Jack Chen, Hongbin Zhang, Patrick Sorgeloos, Amir Sagi, Acacia Alcivar-Warren, Zhanjiang Liu, Lei Wang, Jue Ruan, Ka Hou Chu, Bin Liu, Fuhua Li, and Jianhai Xiang. Penaeid shrimp genome provides insights into benthic adaptation and frequent molting. *Nature Communications*, 10(1), January 2019.

[10] Kristina Mead and Roy Caldwell. Mantis shrimp: Olfactory apparatus and chemosensory behavior. In *Chemical Communication in Crustaceans*, pages 219–238. Springer New York, 2010.