**Title**

Interdisciplinary Approaches to Conservation Biology: From management implications for an endangered coastal species to the effects of active learning in higher education

**Author**

Ha, Benjamin

**Publication Date**

2023

UNIVERSITY OF CALIFORNIA

Los Angeles

# Interdisciplinary Approaches to Conservation Biology: From management implications for an endangered coastal species to the effects of active learning in higher education

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Biology

by

Benjamin Andrew Ha

2023

# ABSTRACT OF THE DISSERTATION

## Interdisciplinary Approaches to Conservation Biology: From management implications for an endangered coastal species to the effects of active learning in higher education

by

Benjamin Andrew Ha

Doctor of Philosophy in Biology

University of California, Los Angeles, 2023

Professor David K. Jacobs, Chair

Contributions to conservation biology research are traditionally approached from an ecology and evolutionary biology perspective, but education research can be equally valuable. Population genetics research can inform management decision-making while education research can improve how we teach conservation. My dissertation tackles conservation biology from an interdisciplinary framework. This involves studying the population genetics of the endangered tidewater goby and investigating drawing as an active learning tool in a Restoration Ecology course. For chapter 1, I assessed how the genetic structures between tidewater goby (Genus: *Eucyclogobius*) populations relate to high rainfall events by analyzing rainfall data and 14 microsatellite loci from 526 samples collected in 1990 and 2008 from the central coast of California. There is evidence that supports tidewater goby dispersal is dependent on heavy rainfall that subsequently breaches lagoons. Results also show that two adjacent management subunits are genetically similar in the 2008 samples compared to 1990, which suggests the subunits should be combined. For chapter 2, I investigated how landscape variables inform genetic structures between tidewater goby populations by analyzing satellite imagery data and 14 microsatellite loci from 1346 individuals covering four management units over an 18-year timespan. Results suggest heterogeneous landscapes that include softer substrates may facilitate movement between sites. There is also evidence that harder substrates, headlands, and kelp vegetation may limit dispersal. For chapter 3, I explored how drawing activities in an ecological restoration course affect learning outcomes and whether students continue to use drawing as a study tool after the intervention. Results suggest students use their notes and lecture slides more frequently than drawing activities to study for the midterm. Students who were instructed to draw (drawers) as part of their response for a midterm scored significantly higher than students who did not (non-drawers). Drawers also had a more complex correlation network of paired works in their midterm responses compared to non-drawers. Collectively, this dissertation expands on how we traditionally address conservation biology by offering insights from population genetics biology combined with innovative applications of pedagogy in higher education.

The dissertation of Benjamin Andrew Ha is approved.

Kyle C. Cavanaugh

Thomas Welch Gillespie

Jonathan Puritz

James W. Stigler

David K. Jacobs, Committee Chair

University of California, Los Angeles

2023

# DEDICATION

I want to dedicate the completion of this PhD to Taylor Brown and Sammy Mensah, two incredible UCLA scholars who did not have the opportunity to complete their PhDs before passing away. I also want to dedicate this to those who identify as first-generation undergraduates, first in their family to pursue a graduate degree, immigrants, children of immigrants and war refugees, and anyone else, especially people of color, who have felt like their presence, knowledge, and resilience have been oppressed during their time in higher education. I see all of you. Let this dissertation serve as motivation and proof that we belong in systems that were not originally built for us. Only now do I understand that anyone can be a researcher. And if you cannot become one, then something is wrong with the system, not you. For all of those who identify as White, may the completion of this PhD inspire you to learn more about how racism relates to your life as an individual, to your career, to your research, and to your mentees, and that you continue to reflect on your knowledge and experience with your White colleagues and grow together. To quote Heather McGhee, an economist and policy advocate, “Racism has a cost for everyone.”

# TABLE OF CONTENTS

List of Tables

List of Figures

List of Supporting Information

# ACKNOWLEDGMENTS

*"Caring for myself is not self-indulgence, it is self-preservation,*
*and that is an act of political warfare." – Audre Lord*

I want to thank my lab mates from the Jacobs Lab for their emotional support and nerdy community throughout my PhD career: Mira Abrecht, Jeana Drake, Hayden Speck, Brenton Spies, Rachel Turba, and Paige Zhang, with special thanks to Elizabeth Heath-Heckman for training and guiding me during the early years of my program. I want to thank my other lab mates from the Teaching and Learning Lab for openly welcoming me into their incredible community and filling my life with laughter: Adam Blake, Caylor Davis, Laura Fries, Karen Givvin, Ji Son, Jim Stigler, Claudia Sutter, Mary Tucker, Alice Xu, and Icy Zhang. I will forever cherish the day we all spent lab meeting on zoom eating extra spicy instant ramen at the same time. Thank you to my undergraduate research mentees for all your contributions to lab and for trusting me to be your mentor: Brendan Soo, Madeline Trotter, and Ellison Yang. Thank you to all my other undergraduate mentees for bringing so much joy to my life and also trusting me to be your mentor: Liesel Davis, Hamilton Oh, and Samuel Rivera. All of you will forever have a special place in my heart. Thank you to my PhD advisor, David Jacobs, and committee members for all the guidance throughout my PhD career, especially Tom Gillespie for all the emotional support during my final year.

Thank you to my EEB community *et al.* with special shoutouts to (in no specific order): Christopher Kyriazis, Shalanda Grier, Christiane Jacquemetton, Aidan Howenstine, Win Hongjamrassilp, Kenji Hayashi, Kaija Gahm, Rachel Kennison, Peggy Fong, Pam Yeh, Paul

Barber, Leryn Gorlitsky, and Marco Morselli. And a very special thanks and gratitude to everyone who helped me before I started the PhD program: José de la Torre, Carmen Domingo, Sarah Cohen, Beth McPartlan, Diana Azurdia, Sonia Zarate, and Dwayne Simmons. I would have never been accepted into, or let alone considered pursuing, a PhD program without any of your collective wisdom, guidance, and support.

As a first-generation undergraduate, first-generation and first in my family to pursue a graduate degree, a child of immigrants and war refugees who escaped the "Vietnam" War (a.k.a. the American War in Vietnam), and a queer Vietnamese American, I would not be here without my family and friends. Em yêu gia đình và các bạn nhiều lắm. Tôi rất hãnh diện là người Việt Nam. Thank you to my parents for their sacrifices, which enabled me to pursue higher education. To my loving and supportive siblings, I am forever grateful to have you in my life. A special shoutout to my dodgeball community for bringing me newfound joy in life and a much-needed escape from grad school during the final two years of my program. And a second thanks to dodgeball for introducing me to my partner, who has been so loving and patient with listening to all of my experiences. Meow! Thank you to everyone else for being part of this PhD journey. I am grateful for every experience, both good and bad, that has led me to this accomplishment.

And for the last time, I would like to acknowledge that funding was provided by the National Science Foundation Graduate Research Fellowship Program (Grant no. 1650604). Thank you to the 2016-2017 review committee, who will forever be anonymous, for believing in my capacity to become a researcher when all other graduate programs did not even believe in me themselves. I was rejected by all the MS and PhD programs I applied to during the 2016-2017 admissions cycle, and I know deep down that I would have never been accepted into a PhD program if I had not been awarded the 2017 NSF GRFP. Over and out beezies!!

For Chapter 1, Benjamin A. Hà, Janet C. Buckner, Jonathan Y. Sim, and Greer A. Dolby completed data collection and analysis. Dan Holland, Camm C. Swift, Brenton T. Spies, and David K. Jacobs conducted fieldwork and collected tidewater goby samples. Janet C. Buckner and Camm C. Swift helped edit the manuscript. Benjamin A. Hà and David K. Jacobs wrote the manuscript.

For Chapter 2, Benjamin A. Hà, Jonathan Y. Sim, Kyle Cavanaugh, and David K. Jacobs completed data collection. Dan Holland, Camm C. Swift, and David K. Jacobs conducted fieldwork and collected tidewater goby samples. Benjamin A. Hà conducted data analysis. Kyle Cavanaugh helped edit the manuscript. Benjamin A. Hà and David K. Jacobs wrote the manuscript.

For Chapter 3, Benjamin A. Hà completed data collection and analysis. Benjamin A. Hà and Leryn Gorlitsky processed IRB paperwork. Leryn Gorlitsky was the instructor for the course in which the interventions and experiments occurred. Karen Givvin, Ji. Y. Son, and James W. Stigler provided guidance on research methods and design. Benjamin A. Hà and James W. Stigler wrote the manuscript.

# BENJAMIN ANDREW HÀ

*Curriculum Vita*

## EDUCATION

**New York University, New York, NY** **2011**

B.A. English and American Literature, honors scholar

## CERTIFICATIONS

**Facilitating Mentor Training,** *Center for the Improvement of Mentored Experiences in Research* **In progress**

Completing hours to become a certified facilitator to lead and develop trainings on effective mentorship for leadership in academia and industry

**Diversity, Equity and Inclusion (DEI) in the Workforce Certificate,** *University of South Florida* **2023**

Completed 14+ hours of certified training for HR professionals to implement DEI in the workforce

**Mentor Training Program,** *Center for the Improvement of Mentored Experiences in Research* **2021**

Completed 15+ hours of certified training on effective mentorship

**Education Scholar,** *Center for the Integration of Research, Teaching, and Learning* **2019**

Completed 45+ hours of certified training on how to implement evidence-based active learning and inclusive pedagogies in higher education

## SELECT WORK EXPERIENCE

**PhD Candidate** **Sep 2017 – present**

**UCLA's Department of Ecology and Evolutionary Biology (EEB), Los Angeles, CA**

- Individually secured funding from the EEB Department and UCLA Life Sciences Division to organize and lead 3 separate graduate program events, including UCLA EEB's first recruitment event for prospective graduate applicants by collaborating with various centers on campus
- Continue to advise on strategic planning and initiatives for the department, such as consulting on faculty professional development and graduate education curriculum development, and developing a climate survey to collect data on graduate student experiences and outcomes
- For 3+ years, managed hiring, training, and mentorship for 3 undergraduate researchers
- For 5+ years, managed lab operations, including administration, undergraduate professional development, and budget management

**Graduate Student Researcher** **Jul 2020 – Apr 2021**

**UCLA's Center for Education Innovation and Learning in the Sciences, Los Angeles, CA**

- Inaugurated UCLA's first campus-wide efforts for trainings on diversity, equity, inclusion, and anti-racism; facilitated trainings for faculty, staff, postdocs, and graduate students across all disciplines
- Provided thought leadership and consultation on training materials and facilitated content equivalent to 50+ total hours of trainings across 40+ total workshops for 1200+ participants

**Administrative Coordinator** **Oct 2014 – Aug 2017**

**Stanford University's Emmett Interdisciplinary Program in Environment and Resources, Palo Alto, CA**

- Supported graduate and professional degree programs (PhD, Dual/Joint MS for JD, MD, and MBA)
- Managed program operations for ~120 total enrolled students, including administration, events management, admissions, recruitment, and strategic planning

**Program Coordinator** **Jan 2012 – Mar 2014**

**UCLA's Undergraduate Research Center in the Sciences, Los Angeles, CA**

- Supported 6 undergraduate academic and research programs, which provide programming to retain minoritized undergraduates in biomedical sciences
- Managed program operations for ~150-200 total students, including administration, events management, admissions, recruitment, hiring, and strategic planning

## TEACHING EXPERIENCE

**Lead Instructor/Teaching Fellow** **Winter 2022**
**EE BIOL 98T: The History and Racism of Biology Research: Working towards an Anti-Racist Science Future**
*University of California, Los Angeles (UCLA), Los Angeles, CA*

- Through a one-year process, developed the course proposal and designed curriculum for an undergraduate seminar on the history and racism of 9 biology-related topics
- Managed all course operations for 17 enrolled undergraduates for 3 hours of lecture per week for 10 weeks, including instruction, administration, and recruitment

Additional UCLA courses:

| | |
|---|---|
| Guest Lecturer. EE BIOL 200B: Evolutionary Biology | Winter 2023 |
| Guest Lecturer. EE BIOL 136: Ecological Restoration | Winter 2023 |
| Guest Lecturer. EE BIOL 136: Ecological Restoration | Spring 2022 |
| Teaching Fellow. EE BIOL 97XAB: PEERS Program Seminar | Fall 2021 |
| Teaching Associate. LS 110: Career Exploration in Life Sciences | Spring 2021 |
| Teaching Associate. LS 495: Preparation for College-Level Teaching in Life Sciences | Fall 2020 |
| Guest Lecturer. EE BIOL 136: Ecological Restoration | Spring 2020 |
| Teaching Associate. EE BIOL 106 + 124a: Marine Biology Field Quarter | Fall 2019 |

## PRESENTATIONS

**Not listed: presented a poster at 7 conferences and 1 invited seminar talk*

**Oral. Hà, B.A.**, Buckner, J.C., Sim, J., Dolby, G.A., Holland, D., Swift, C.C., Spies, B.T., Jacobs, D.K. *Hydrologic Control of Metapopulation Dispersal in the Endangered Tidewater goby (Eucyclogobius newberryi) on the Central California Coast*. 2020 Society for the Advancement of Chicanos/Hispanics and Native Americans in Science (SACNAS) Conference, virtual

## PUBLICATIONS

**Not listed: 3 co-author publications and 1 co-author manuscript in prep*

**Hà, B.A.**, Foxx, K., Mensah, S.T., Barber, P.H., Kennison, R.L. (2023) Interdisciplinary approaches to advancing anti-racist pedagogies in ecology, evolution, and conservation biology. *Trends in Ecology and Evolution*

## COMMUNITY OUTREACH

**Not listed: participation on 3 committees, 9 panels, and 2 mentoring programs*

**UCLA EEB Anti-Racism Task Force** **2020 – 2021**

- Developed strategic planning to advance practices and policies to promote diversity, equity, inclusion, and anti-racism for graduate students and faculty
- Topics included curriculum, professional development, recruitment, admissions, and retention

## FELLOWSHIPS, GRANTS, AND AWARDS

**Not listed: 15 additional fellowships, grants, and awards*

| | |
|---|---|
| 2022 – 2023 | UCLA Dissertation Year Fellowship, $38000 |
| 2017 – 2022 | National Science Foundation Graduate Research Fellowship Program, $138000 |
| 2021 – 2022 | UCLA Collegium of University Teaching Fellowship |
| 2022 | UCLA Life Sciences Division Excellence Award in Educational Innovation |
| 2021 | UCLA Life Sciences Division Excellence Award in the Promotion of Diversity and Inclusion |
| 2021 | Equity, Diversity, and Inclusion Service Award in Ecology and Evolutionary Biology |
| 2020 | Best Graduate Oral Presentation in Ecology and Evolutionary Biology, SACNAS Conference |
| 2020 | La Kretz Center and Stunt Ranch Research Grant, $7000 |
| 2019 | UCLA Affiliates Scholarship, $10000 |

# Chapter 1

# Hydrologic Control of Metapopulation Dispersal in the Endangered Northern Tidewater Goby (*Eucyclogobius newberryi*) on the California Coast

Benjamin A. Hà,[1] Janet C. Buckner,[1,2] Jonathan Y. Sim,[1] Greer A. Dolby,[1,3] Dan Holland,[1] Camm C. Swift,[4] Brenton T. Spies,[1,5] David K. Jacobs[1]

[1] Department of Ecology and Evolutionary Biology, University of California, Los Angeles
[2] Department of Biology, University of Texas at Arlington
[3] Department of Biology, University of Alabama at Birmingham
[4] Natural History Museum of Los Angeles County
[5] Environmental Science and Resource Management Program, California State University, Channel Islands

## Abstract

California coastal lagoons function in response to seasonal as well as episodic precipitation and streamflow events. During much of the year, these lagoons provide some of the only freshwater habitats available regionally, but during heavy stream flow events, lagoons open to the sea providing access for an array of organisms, such as marine and anadromous species. Consequently, this hydrologically variable environment supports a range of species of special concern. Here, we link the metapopulation genetics of the federally endangered lagoon resident, the Northern Tidewater goby (*Eucyclogobius newberryi*), to variable stream flow and lagoon opening. Analysis of microsatellites, following a dry period (1990) and a wet period (2008), shows that dispersal is facilitated in the wet year relative to the dry year as assessed by an assignment test and other population genetic measures. Thus, we link the opening of lagoons to *E. newberryi* dispersal and metapopulation processes, confirming the long-suspected association between the hydrologic breaching of lagoons and marine dispersal as adults. Dispersal, when it transpires, is also mostly along sandy instead of rocky coasts. In the Central Coast management unit, we infer one of the management subunits, CC1, was extirpated and recolonized from the adjacent management subunit, CC2, and therefore, argue that CC1 and CC2 should be consolidated. This work demonstrates the connection between varying precipitation and the metapopulation processes of this endangered species, a potential model organism that may be used to interrogate the relation of the connectivity of these dynamic wetland habitats that support a suite of at-risk species. California coastal lagoons experience a suite of anthropogenic impacts that include direct as well as climate-induced anthropogenic effects. Climate change will profoundly affect these and other Mediterranean climate-dependent habitats. These results provide a model for variable hydrologically driven systems at risk from climate change.

## Introduction

Coastal lagoons form at stream mouths all along the California coast where coastal bar, beach, and dune formations isolate the lagoon from the ocean during periods of low streamflow (Behrens et al., 2013; Jacobs et al., 2011). While isolated from the sea, lagoons typically form brackish water or freshwater systems. In addition to storms and high wave energy, stream flow during periods of rainfall can open or breach lagoons, causing them to have a tidal and more marine influence of salinity and temperature from the ocean (Hoeksema et al., 2018; Rich & Keller, 2013; Williams, 2014). Lagoon waters can lower during drought and can often rise above the high tide level, perched behind the height of the beach berm (Jacobs et al., 2011; Williams, 2014). Thus, lagoon dynamics are thought to depend significantly on rainfall and the resultant streamflow.

While the breaching behavior of lagoons can be modeled in relation to streamflow (Kraus, 2003; Rich & Keller, 2013), stream hydrology is the primary driver of lagoon breaching (Orescanin & Schooler, 2018; Schooler, 2017). It is important to note that lagoons can experience a variety of intermediate closure states (Jacobs et al., 2011). In addition, more extreme hydrologic processes can be important where very heavy rainfall/stream flooding can scour or rejuvenate lagoonal habitats (Ranasinghe & Pattiaratchi, 2003). In contrast, drought and low-flow conditions can lead to desiccation of lagoons or infilling with sediment (Lafferty et al., 1999a). The physical environments of a lagoon vary in response to hydrology on timescales from intra-annual to decadal, which can greatly modify the condition of habitats and the potential connectivity of their inhabitants through time.

Hydrologically-mediated variation in habitats through time has numerous biological consequences for a range of species that take advantage of the local freshwater resources.

California lagoons support a range of species of special concern including red-legged frogs (*Rana aurora draytonii*), two-striped garter snakes (*Thamnophis hammondii*), western pond turtles (*Clemmys marmorata*), and steelhead trout (*Oncorhynchus mykiss*) (Bond et al., 2008; Hayes et al., 2008; Huber & Carlson, 2020; Lafferty & Page, 1997; Reese & Welsh Jr, 1998; Reis, 1999). However, these taxa can move to or utilize other freshwater resources and are capable of a variety of modes of dispersal. In contrast, the tidewater gobies, both the Northern Tidewater goby (*Eucyclogobius newberryi*) treated here and the Southern Tidewater goby (*Eucyclogobius kristinae*), are lagoon specialists that are rarely found far upstream and have never been observed in the open ocean as adults or larvae (Swift et al., 1989). Consequently, the metapopulation dynamics and genetic divergence of tidewater gobies should be dominantly influenced by the ephemeral and variable nature of individual lagoons, which, in turn, makes dispersal between lagoons largely dependent on high rainfall events that breach more than one lagoon simultaneously. Consequently, the comparative population genetic reconstruction of migration in wet and dry times provides an effective tool to examine the relation between lagoon hydrologic processes and coastal marine dispersal in the tidewater gobies.

Tidewater gobies exhibit metapopulation dynamics that vary from north to south along the California coast (Earl et al., 2010; Lafferty et al., 1999a, 1999b). Populations are genetically isolated on local scales and deeply subdivided between regions (Dawson et al., 2001, 2002; Earl et al., 2010). Spatial genetic structure informed designations between and within six management units for the Northern and Southern Tidewater goby (Swift et al., 2016; U.S. Fish and Wildlife Services, 2005, 2013, 2014). Each management unit consists of one or more subunits, each of which represents a distinct metapopulation (Figure 1.1). Population genetic studies suggest that the tidewater goby is one of the most genetically subdivided vertebrates along the Pacific Coast of California (Dawson et al., 2001; Earl et al., 2010; McCraney et al., 2010; Mendonca et al., 2001). Large geographic breaks in distance and rocky headlands between lagoons can limit gene flow (Dawson et al., 2001; U.S. Fish and Wildlife Services, 2005), such as Piedras Blancas in the north and Point Estero in the south of the Central Coast management unit (CCU; Figure 1.1). They may serve as geographical and presumptive genetic breaks between tidewater goby subunits. Piedras Blancas separates subunit CC1 and subunit CC2, and Point Estero separates subunit CC2 and subunit CC3.

Dispersal between coastal lagoons requires longshore ocean movement between simultaneous breached lagoons (Lafferty et al., 1999a, 1999b; Swenson, 1999). Populations in the northern management unit are often isolated and known to be persistent (McCraney et al., 2010) while source-sink and local metapopulation processes are evident further south (Earl et al., 2010; Swift et al., 2016). Thus, the metapopulation behavior of tidewater gobies appears to reflect the latitudinal hydrologic conditions across management units, where the north has more consistent annual rainfall and more persistent populations whereas the south has more extreme hydrologic variation and a higher frequency of extirpation. The extirpation of tidewater goby populations has often been associated with drought and the desiccation of lagoons; although, artificial breaching may also cause extirpation (Lafferty et al., 1999a; Swift et al., 2018) or a severe reduction in population size (U.S. Fish and Wildlife Services, 2005). Despite these general inferences on extirpation and recolonization based on presence/absence data, the potential relation between dispersal and hydrologic events has only been indirectly assessed.

Here, we employ a temporal analysis of microsatellite and hydrological data comparing Northern Tidewater goby (hereafter, used interchangeably with tidewater gobies) samples after dry (1990) and wet (2008) periods from sites across the CCU. This allows the assessment of the relation between tidewater goby dispersal and rainfall events that breach lagoons. We find that movement between sites occurs in association with the high rainfall episodes, presumably because these events resulted in the simultaneous breaching of lagoons. Substantial dispersal is not evident following an extensive period of low rainfall. This time period likely lacked stream flow and coordinated breaching sufficient to support the conditions required for the dispersal of this species. We additionally find that dispersal predominantly occurs along sandy coastlines. Lastly, we note that the data reveal an extirpation and recolonization event between the two years studied, such that the populations are currently no longer distinct. As a result, the data suggests subunit CC1 in the CCU is no longer a distinct metapopulation from subunit CC2, and in turn, may no longer merit separate subunit status.

## Methods

### *Sample sites and collections*

A total of 526 tidewater goby individuals were collected by seine net (15x6 feet with 8 inch mesh) from 13 sites covering all three tidewater goby subunits within the CCU in two separate years (Table 1.1). From July through August 1990, 136 individuals were collected from 10 different sites (Arroyo del Corral, COR; Arroyo Laguna, LAG; Arroyo de Tortuga, TOR; Arroyo del Puerto, PUE; Little Pico Creek, LPC; Pico Creek, PIC; San Simeon Creek, SIM; Villa Creek, VIL; San Geronimo Creek, GER; Cayucos Creek, CAY). In February 2008, 390 individuals were collected from 13 sites (the 10 sites mentioned above, plus Broken Bridge Creek, BRO; Santa Rosa Creek, ROS; Little Cayucos Creek, LCA). Three sites (BRO, ROS, and LCA) were only sampled in 2008. The sample sizes were also smaller and less uniform in 1990 than in 2008 given that the California Department of Fish and Wildlife (previously known as

California Department of Fish and Game) limited the number of gobies collected to 20 individuals per site.

Samples were collected by Dan Holland in 1990 prior to the federal listing of tidewater gobies under the Endangered Species Act with permission from the California Department of Fish and Game. Samples collected in 2008 were collected under a federal permit to Camm Swift from the California Department of Fish and Game (Permit # SCP-2679) and U. S. Fish and Wildlife Service (Permit #TE-793644).

The monthly rainfall in inches for this region was collected by a rain gauge at the Morro Bay Fire Department in California (Western Regional Climate Center).

### *DNA extraction and genotyping*

Tidewater gobies were preserved on ice in 95% ethanol or liquid nitrogen in the field prior to being stored at -80℃ until DNA extraction. Whole genomic DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit following the manufacturer's instructions. Tidewater gobies were genotyped at 14 microsatellite loci (*ENE2* through *ENE21*) (Appendix S1.1, Supporting Information). *ENE2* and *ENE3* were based on Mendonca et al. 2001 and *ENE5* through *ENE21* were based on Earl et al. 2010. Genotypes were established by using Qiagen Multiplex PCR kits with a fluorescent dye-labeled M13F (-20) primer (Boutin-Ganache et al., 2001). Forward primers included M13 "tails" for labeling and reverse primers were unlabeled. Samples were prepared according to the manufacturer's guidelines. PCR was performed in a thermocycler using the following profile: 95°C for 15 min; 25 cycles of 94°C for 30 sec, 55°C for 90 sec, 72°C for 60 sec; then 21 cycles of 94°C for 30 sec, 50°C for 90 sec, 72°C for 60 sec, with a final extension at 60°C for 30 min. Products were run on an ABI 3730KL capillary sequencer and scored using ABI GENEMAPPER version 3.0 (Applied Biosystems).

### *Data analysis*

STRUCTURE 2.3 (Pritchard et al., 2010) was used to estimate the number of population subdivisions ($K$) and infer population structure using Bayesian clustering from microsatellite data. We employed the admixture model with 20 replicates for each $K$ number of clusters permitted in the analysis between 2 and 40 using a burn-in of 100,000 iterations followed by 100,000 Markov-Chain Monte-Carlo (MCMC) steps. STRUCTURE HARVESTER (http://taylor0.biology.ucla.edu/struct_harvest) was used to assess and visualize the likelihood values across multiple values of $K$ to identify the number of genetic groups that best fit the data estimated by the optimum $K$ value (Earl & vonHoldt, 2012).

GenAlEx 6.5 (Peakall & Smouse, 2012) was used to calculate the expected heterozygosity, pairwise $F_{ST}$, and average number of alleles per locus. Expected heterozygosity ($H_{exp}$) is the probability an individual will be heterozygous at a given locus. The $H_{exp}$ values described in this study are the average expected heterozygosity across all 14 microsatellite loci. Pairwise $F_{ST}$ assesses genetic differences between pairs of populations. Arlequin 3.0 (Excoffier et al., 2005) was used to calculate the Garza-Williamson index (G-W). The G-W index (Garza & Williamson, 2001) is based on the ratio of the number of microsatellite alleles to the range in allele size to detect genetic bottlenecks. Graph generation and additional statistical analyses (z-score, Wilcoxon rank sum test, and t-test) were performed in R (R Core Team, 2013). The Wilcoxon rank sum test was calculated for group comparisons using $H_{exp}$ and the average number of alleles per locus since the data was not normally distributed. T-tests were calculated for pairwise $F_{ST}$ and the G-W index.

We used assignment tests generated in Arlequin 3.5 (Excoffier & Lischer, 2010) conducted on the whole dataset to generate an inference of the population source of all sampled individuals. The genotypic "assignment" of individuals to sites other than where they were not collected provides a maximum likelihood probability assessment of directionality and the number of migrating fish in the sample that dispersed between specific sites prior to being collected. Given that the results are likely influenced by the different number of individuals collected from each year (i.e., samples collected in 1990 were significantly smaller than samples collected in 2008), we conducted a rarefaction resampling exercise to provide a statistical confirmation of the expected number of individuals that migrated between sites in 1990. For resampling, we generated 1000 subsets by randomly drawing microsatellite data from the larger 2008 data set equaling the number of individuals at each of the 10 sites in 1990 that were also sampled in 2008. We then performed assignment tests on each subset using Arlequin 3.5 as described above. Finally, we calculated the percentage of 2008 subsets that exceeded the number of migrants recovered in 1990 as a measure of the probability that differences were due to chance.

## Results

### *Dispersal associated with high rainfall and subsequent lagoonal breaching*

Assignment tests attribute most individuals to the site from which they were collected; however, two individuals were assigned to a different lagoon from its original collection site in 1990 while 41 individuals were assigned to a different lagoon from its original collection site in

2008 (Figure 1.2). The assignment tests strongly support that there was a higher migration rate in 2008 than in 1990. A rarefaction resampling analysis confirmed that these differences in the number of estimated migrants is significantly lower in 1990 and did not result from differences in sample size between the two years (Appendix S1.2, Supporting Information). The higher migration rate among the 2008 samples is consistent with significantly lower $F_{ST}$ values between adjacent pairwise sites in 2008 compared to 1990 (paired t-test; p-value < 0.05; Appendix S1.3, Supporting Information). Additionally, expected heterozygosity in the 2008 samples is higher than the samples collected in 1990 (Figure 1.3). This difference is statistically significant when including Arroyo del Corral (hereafter, referred to as Corral; paired Wilcoxon rank sum test, p-value < 0.05) and becomes more statistically significant when excluding Corral (paired Wilcoxon rank sum test, p-value < 0.0001). There is also a significantly higher average number of alleles per locus in the 2008 samples compared to the 1990 samples regardless of including or excluding Corral (paired Wilcoxon rank sum test; p-value < 0.0001; Figure 1.3).

Our collections in February 2008 followed seven inches of rainfall in January as reported by the Morro Bay Fire Department (Western Regional Climate Center; see Figure 1.4). Field observations (i.e., visual examination of breached sand bars) during collection in February 2008 indicated that all systems sampled had recently breached following this January precipitation event. Only one month within the 90 months prior to each collection date reached or exceeded a total of 7 inches of monthly precipitation: one month prior to the February 2008 collection and 90 months prior to the August 1990 collection.

To further explore the rainfall distribution, we examined the data using a threshold value of ≥4 inches of rainfall per month as an assumption for generating breaching events (Appendix S1.4, Supporting Information). With this threshold, the most recent event prior to the February

2008 sampling is one month compared to the 20 months prior to the August 1990 sampling. Furthermore, we only found 11 total such events 90 months prior to the February 2008 collection compared to the 6 total events 90 months prior to the August 1990 collection. There was also a longer time lag between months with ≥4 inches of rainfall in 1990 compared to 2008.

### *Large geographic break limits gene flow between subunits*

The assignment tests do not show movement across the northern rocky headland, Piedras Blancas, nor the southern rocky headland, Point Estero; migration was only inferred via the assignment tests across sandy coastlines within subunits (Figure 1.2). Furthermore, at all $K$ values, STRUCTURE analyses show strong evidence of genetic separation between the two sites separated by Point Estero, regardless of collection year (Figure 1.5). $F_{ST}$ values are also higher across this large geographic break (In 1990, $F_{ST} = 0.057$ between San Simeon Creek and Villa Creek; in 2008, $F_{ST} = 0.083$ between San Simeon Creek and Villa Creek; see Appendix S1.3, Supporting Information) than between adjacent sites on either side of the Point Estero geographic break (In 1990, $F_{ST} = 0.032$ between Pico Creek and San Simeon Creek and $F_{ST} = 0.039$ between Villa Creek and San Geronimo Creek; in 2008, $F_{ST} = 0.018$ between Pico Creek and San Simeon Creek and $F_{ST} = 0.057$ between Villa Creek and San Geronimo Creek). There seems to be a consistent geographic barrier at Point Estero; however, the barrier south of Corral at Piedras Blancas is less consistent as described below.

### *Evidence for extirpation of Corral followed by dispersal and recolonization from geographically nearby sites*

STRUCTURE analyses ($K$>2) in 1990 shows Corral to be genotypically different from the two nearest sites to the south, Arroyo Laguna (hereafter, referred to as Laguna) and Arroyo de Tortuga (hereafter, referred to as Tortuga); however, in 2008, Corral is genotypically similar to Laguna and Tortuga (Figure 1.5). In the 1990 samples, the pairwise $F_{ST}$ values are higher between Corral and Laguna (0.077) and Corral and Tortuga (0.117), whereas in 2008, the pairwise $F_{ST}$ values are lower between Corral and Laguna (0.040) and Corral and Tortuga (0.041) (Appendix S1.3, Supporting Information). Additionally, the pairwise $F_{ST}$ value between the 1990 and 2008 sample years at Corral is significantly higher than any other within-site interannual pairwise $F_{ST}$ comparison (z-score: 2.666, p-value = 0.007; Appendix S1.5, Supporting Information). In contrast, the assignment tests do not recover movement of tidewater gobies to or from Corral with management subunits to the south (Figure 1.2).

Additionally, the 1990 samples from Corral are significantly higher in expected heterozygosity than those from 2008 (paired Wilcoxon rank sum test; p-value < 0.005; Figure 1.3). The 2008 Corral samples also have a lower average number of alleles per locus compared to the 1990 Corral samples; although, this is not statistically significant (paired Wilcoxon rank sum test; p-value = 0.0708; Figure 1.3). The G-W index, a proxy for genetic bottleneck, is higher in the 2008 Corral samples relative to the 1990 samples (Appendix S1.6, Supporting Information).

## Discussion:

### *Dispersal associated with high rainfall and subsequent lagoonal breaching*

Stream hydrology is thought to be the primary driver of lagoon breaching (Jacobs et al., 2011; Orescanin & Schooler, 2018; Schooler, 2017). It appears that 7 total inches of monthly rainfall is enough to breach lagoons based on field observations during the February 2008 collection. While we assumed a lower threshold of ≥4 inches of rainfall per month is sufficient to breach some lagoons, it may not be sufficient to breach all lagoons. This lower criterion for rainfall to produce breaching events may further depend on the amount of precipitation proceeding over consecutive months.

We also observed a longer lag period of <4 inches of monthly rainfall within the 90 months prior to the 1990 field collection compared to that in 2008. This dry period prior to the 1990 sampling is consistent with the relative absence of El Niño between two large El Niño events of 1982-1983 and 1997-1998 (National Weather Service Climate Prediction Center). Whereas prior to the 2008 sampling, El Niño events have been observed in 2002-2003, 2004-2005 and 2006-2007, including an El Niño event that developed in September 2007 and peaked in February 2008 (National Weather Service Climate Prediction Center). Given that breaching is thought to be critical to tidewater goby dispersal, rainfall prior to 2008 is expected to promote more movement than during the low rainfall periods prior to 1990.

Our samples collected in February 2008 following a month with 7 inches of rainfall appear to have captured a relatively large number of migrants as determined by the assignment tests. This is consistent with the pairwise $F_{ST}$ values being significantly lower between adjacent sites in 2008 than those in 1990, which suggests greater genetic similarity due to increased movement between sites. Contrastingly, the pairwise $F_{ST}$ values are significantly higher between adjacent sites in 1990, which suggests greater genetic differentiation in the absence of migration from the assignment tests prior to the 1990 sampling. This difference in migration between years with different preceding rainfall histories supports the role of rainfall, streamflow, and extensive lagoon breaching in the dispersal of tidewater gobies. The significantly higher expected heterozygosity and allelic diversity in the 2008 samples for all sites is also likely a product of higher migration in 2008 and a period of isolation and local bottlenecking prior to 1990.

The timing of our samples collected the month immediately following a heavy rainfall event could have potentially detected movement better as opposed to sampling ≥6 months afterwards. Tidewater gobies have about a one-year lifespan and can reproduce multiple times during the year with distinct peaks in spawning, often in early spring and later summer (Swenson, 1999). This suggests the timing of sample collection relative to rainfall may play an important role to detect genetic evidence for tidewater goby dispersal. Further work is merited to establish the causal details regarding the precise conditions of hydrologic processes to generate lagoon breaching and, in turn, tidewater goby dispersal.

### *Movement mostly across sandy coastlines with a large geographic break limiting gene flow between subunits*

The CCU is subdivided into three management subunits, which were determined by prior genetic subdivision studies (Dawson et al., 2001; Earl et al., 2010) and the premise that rocky headlands serve as barriers to gene flow (Dawson et al., 2001; U.S. Fish and Wildlife Services, 2005). Assignment tests support this with dispersal detected across more sandy coastlines and no migration across either the Piedras Blancas or Point Estero rocky headlands, which suggests rocky substrates and/or large headlands limit tidewater goby dispersal. This is consistent with previous arguments that dispersal primarily occurs within subunits rather than across subunits separated by rocky boundaries (Swenson, 1999).

This also supports the idea that rocky coastlines and headlands potentially act as dispersal barriers between more interactive metapopulations separated exclusively by sandy substrate. It was hypothesized that tidewater gobies would not enter rocky areas because their light color as a bottom-dwelling fish would subject them to predation in those settings (Swenson, 1997). We observe no assignments indicating migration across these rocky geographic barriers. The distances inferred for migration by the assignment tests also suggest that movements of 5-10 kilometers over sandy substrate are possible. This is a larger value than previously reported based on extirpation and recolonization (Lafferty et al., 1999a).

Contrastingly, closer examination of STRUCTURE plots and pairwise $F_{ST}$ values between the 1990 and 2008 samples at Corral suggest movement across the northern rocky headland is possible, but relatively rare. Further detail is described below.

### *Extirpation of Corral in subunit CC1 and recolonization from geographically nearby sites in subunit CC2*

STRUCTURE analyses and pairwise $F_{ST}$ values support extirpation of Corral followed by dispersal and recolonization from Laguna or Tortuga across the northern rocky headland into Corral sometime between 1990 and 2008. In 1990, pairwise $F_{ST}$ values show Corral is genotypically different from Laguna and Tortuga, which aligns with previous research demonstrating that tidewater goby subunits are genetically subdivided (Dawson et al., 2001, 2002; Earl et al., 2010; McCraney et al., 2010). This also reflects how the management subunits are currently partitioned where the northernmost site, Corral, is part of the CC1 subunit, while Laguna and Tortuga are part of the CC2 subunit (U.S. Fish and Wildlife Services, 2005, 2013); however, the temporal comparison from this study suggests a genetic expansion from the management subunit CC2 to CC1. This is of particular interest since Corral is currently recognized as a separate subunit in the recovery plan (U.S. Fish and Wildlife Services, 2005).

The expected heterozygosity for Corral in 1990 is significantly higher than Corral in 2008, and the STRUCTURE analyses show a completely different genotype group in the two different sample years at Corral, which supports genetically distinct ancestry. The genotype present in the 1990 Corral sample is dramatically different from the geographically nearest sites in subunit CC2, Laguna and Tortuga. This is not the case in 2008. There is no evidence for the genetic remnants from the 1990 Corral samples present in the 2008 samples. Additionally, the most common genotypic class for Corral in 2008 is now uniformly dominant in the individuals sampled. This strongly supports that one or more extirpation events, subsequently followed by recolonization as the mechanism of genetic replacement, occurred sometime during the 18 years between the 1990 and 2008 collections. This is consistent with the expectation of the frequent loss of tidewater goby populations in small lagoons, such as Corral via desiccation (Appendix S1.7, Supporting Information) or other hydrologic events (Lafferty et al., 1999a; U.S. Fish and Wildlife Services, 2005, 2013, 2014).

Furthermore, the genotypic uniformity for Corral in the 2008 samples, in addition to the 2008 Corral samples having lower values of population genetic variation than the 1990 Corral samples, is consistent with a bottleneck following a founder event sourced by a nearby site in the northern region of subunit CC2. The significant difference in expected heterozygosity between collection years further provides evidence for this bottleneck. While the G-W index can be used as a proxy for genetic bottleneck, we suspect the G-W index is less effective than other measures of genetic variation in identifying bottlenecks within the context of this study due to the extensive genetic subdivision and complex history of tidewater gobies. Evidence for bottlenecks is not uncommon though, given that they have been detected in tidewater goby populations in Northern California (Mendonca et al., 2001).

Various anthropogenic and natural factors may have also contributed to the extirpation at Corral, such as habitat degradation, channelization, or desiccation (Lafferty et al., 1996, 1999a; Swift et al., 1993; U.S. Fish and Wildlife Services, 2005). While our data show higher rainfall prior to the 2008 collection, flooding alone is unlikely to have led to extirpation (Lafferty et al., 1999b). While evidence for other recolonization events has been reported in the North Coast Unit, Greater Bay Area Unit, Conception Unit, and LA/Ventura Unit (Kinziger et al., 2015; Lafferty et al., 1999a, 1999b), tidewater goby extinction-recolonization processes are thought to be infrequent (Kinziger et al., 2015; Lafferty et al., 1999a). Therefore, we anticipate that the tidewater gobies inhabiting Corral are unlikely to change genetically in the near future and in turn, recommend U.S. Fish and Wildlife Services to consolidate subunits CC1 and CC2.

### *Tidewater gobies and the changing hydrology and climate in California*

California lagoon breaching driven by rainfall is a product of climate that varies both geographically and temporally, which suggests geographic and temporal variation in tidewater goby metapopulation dynamics and persistence (Lafferty et al., 1999a, 1999b). The tidewater goby genus *Eucyclogobius* ranges across all the coastal counties of California where rainfall and streamflow are greater and far more regular on an annual basis in the north and more episodic in the south. As rainfall events appear to govern dispersal through the breaching of lagoons, climate and climate change should have a strong regional impact on the metapopulation dynamics across the California coastline. In this context, climate variation, in combination with anthropogenic habitat loss, places the Southern Tidewater goby, *E. kristinae,* at greater risk where the species is reduced in range and highly endangered (Swift et al., 2016). It is now restricted to a handful of lagoons on the relatively less-developed landscape of Marine Corps Base Camp Pendleton, where all the habitats have been subjected to extirpation via hydrologic processes and/or anthropogenic effects. The Northern Tidewater goby, *E. newberryi,* is subdivided into five management units where the LA/Ventura Unit appears most at risk, with hydrologic variation playing a significant role.

Climate change impacts on hydrology may be more critical and immediate than changes in sea level. Variation in hydrology may operate on tidewater goby persistence in two ways. Temperatures are likely to increase (Thuiller, 2007) leading to more frequent desiccation, especially in small lagoon habitats, and in turn tidewater goby extirpation. Precipitation is also likely to become more irregular (Berg & Hall, 2015). It is tempting to view these in a purely latitudinal context with warming and more episodic conditions expanding northward starting from the South Coast Unit, where they are already endangering *E. kristinae*, to the LA/Ventura Unit, where subsequently more units inhabited by *E. newberryi* would be impacted. However, this is likely an oversimplification since larger lagoonal systems that frequently open to the sea tend to be cooler at the same latitude and less prone to desiccation (Spies & Steele, 2016). Moreover, upwelling regions associated with Peninsulas (e.g. Point Conception in the Conception Unit) are colder whereas embayment regions (e.g. Monterey Bay in the Greater Bay Area Unit) are warmer than the surrounding coasts (Spies & Steele, 2016).

While California coastal temperatures depend significantly on latitude, other factors involving embayment, upwelling, lagoon size, breaching frequency, etc. have significant impacts at local and regional scales (Spies & Steele, 2016). Climate change in these systems is not expected to be a linear latitudinal phenomenon. For example, a hiatus of California's current upwelling system, which has happened in the past, would have a dramatic effect on temperatures and desiccation of lagoons. El Niño can play multiple roles in these systems, including having negative effects on upwelling and the generation of anomalous rainfall events across much of California (García-Reyes & Largier, 2012; Huyer et al., 2002). In fact, with the work presented here, the 2004-2005 El Niño is correlated with high rainfall events in Central California that likely enhanced dispersal across the study region prior to the February 2008 sample collection. Furthermore, the minimal rainfall for multiple years prior to 1990 is consistent with the relative absence of El Niño between the large El Niño events of 1982-1983 and 1997-1998 (National Weather Service Climate Prediction Center).

The details of how climate changes in the future in terms of increased episodic rainfall combined with warming are likely to have direct effects on the tidewater goby metapopulation process. There are additional indirect effects of warming including the prevalence of warm water species as invasive predators and competitors of tidewater gobies (U.S. Fish and Wildlife Services, 2005). Climatic changes will likely interact with other anthropogenic changes, such as channelization of coastal streams and lagoons. Channelization combined with higher peak rainfall may decrease tidewater goby persistence in the high flow events anticipated with climate change (Swift et al., 2018; U.S. Fish and Wildlife Services, 2005). Thus, this work linking dispersal and precipitation in tidewater gobies demonstrates the need to assess future climate impacts on tidewater gobies, and on the range of other taxa of concern that benefit from these systems.

Table 1.1: Information regarding sample collections, including number of individuals, average number of alleles per locus, and expected heterozygosity ($H_{exp}$) for each respective site and collection year. Sites are listed in geographic order from north to south.

| | | | | **1990 Samples** | | | **2008 Samples** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Site name** | **Site name abbreviation** | **Latitude** | **Longitude** | **No. of individuals** | **Avg no. of alleles per locus** | $H_{exp}$ | **No. of individuals** | **Avg no. of alleles per locus** | $H_{exp}$ |
| Arroyo del Corral | COR | 35°41'05 | 121°17'10 | 12 | 3.36 | 0.470 | 30 | 2.86 | 0.331 |
| Arroyo Laguna | LAG | 35°39'48 | 121°12'39 | 14 | 3.71 | 0.450 | 30 | 4.21 | 0.421 |
| Arroyo de Tortuga | TOR | 34°42'43 | 120°35'50 | 12 | 2.50 | 0.314 | 30 | 3.79 | 0.402 |
| Arroyo del Puerto | PUE | 35°38'35 | 121°11'15 | 12 | 2.57 | 0.354 | 30 | 4.07 | 0.418 |
| Broken Bridge Creek | BRO | 35°38'31 | 121°10'54 | NA | NA | NA | 30 | 3.64 | 0.430 |
| Little Pico Creek | LPC | 35°38'02 | 121°09'46 | 12 | 3.21 | 0.415 | 30 | 4.21 | 0.451 |
| Pico Creek | PIC | 35°36'56 | 121°08'53 | 12 | 3.50 | 0.416 | 30 | 3.93 | 0.454 |
| San Simeon Creek | SIM | 35°35'43 | 121°07'35 | 17 | 4.00 | 0.469 | 30 | 4.43 | 0.479 |
| Santa Rosa Creek | ROS | 35°34'08 | 121°06'36 | NA | NA | NA | 30 | 3.71 | 0.535 |
| Villa Creek | VIL | 35°32'12 | 120°53'31 | 18 | 4.14 | 0.490 | 30 | 4.36 | 0.503 |
| San Geronimo Creek | GER | 35°26'59 | 120°54'27 | 13 | 3.00 | 0.416 | 30 | 4.36 | 0.467 |
| Cayucos Creek | CAY | 35°26'58 | 120°54'26 | 14 | 3.64 | 0.452 | 30 | 4.79 | 0.478 |
| Little Cayucos Creek | LCA | 35°26'51 | 120°54'09 | NA | NA | NA | 30 | 4.36 | 0.461 |

Figure 1.1: Modified figure from U.S. Fish and Wildlife Services, 2005 (left California state map) shows current tidewater goby management unit and subunit designations. California coast map (right) shows the sample collection site's geographic location (black arrows), abbreviated name, sample collection year, and subunits within the Central Coast management unit (CCU). Each subunit represents a metapopulation. Piedras Blancas is the northern rocky headland separating subunits CC1 and CC2 while Point Estero is the southern rocky headland separating subunits CC2 and CC3. Three sites (BRO: Broken Bridge Creek; ROS: Santa Rosa Creek; LCA: Little Cayucos Creek) were only sampled in 2008.

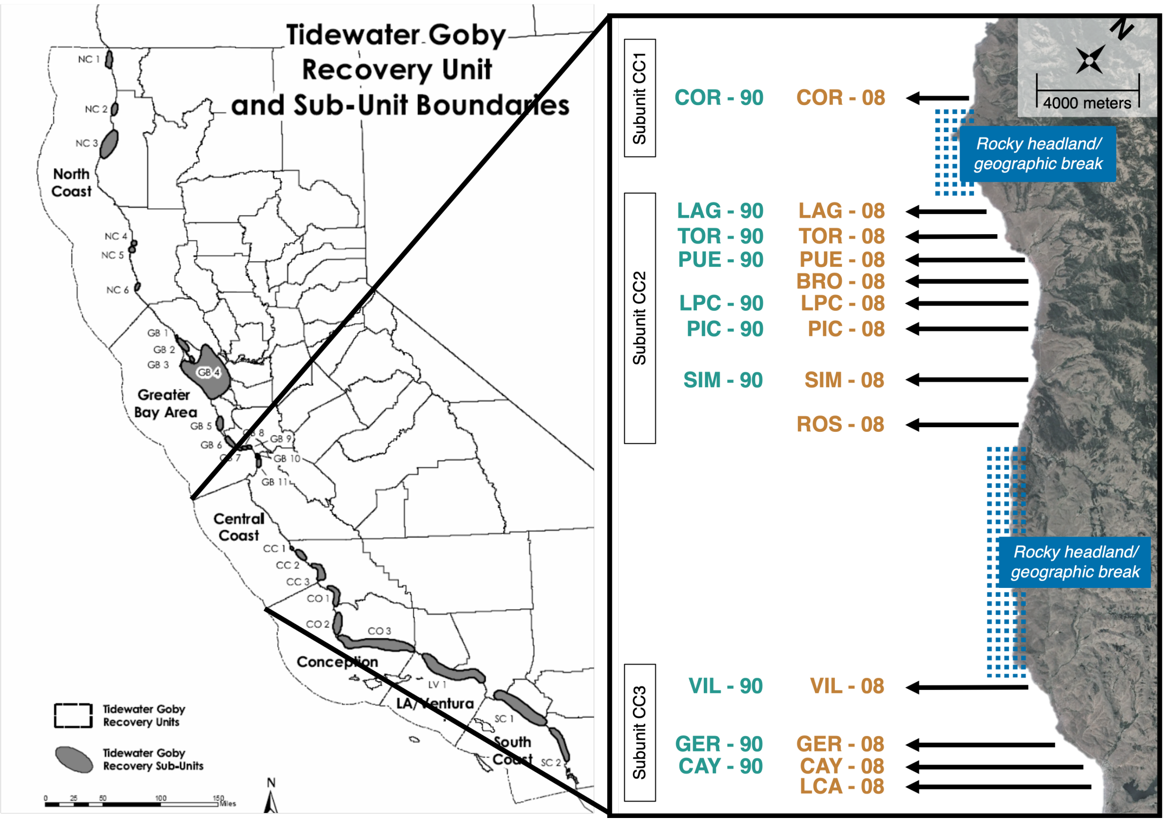


Figure 1.2 Assignment tests assess the stasis or movement of individuals as determined by the highest likelihood assigned for each genotyped individual. Geographic location of collection sites is shown along the California coast (black arrows). Most individuals were assigned to the site from which they were collected (not shown). The quantity and predicted source of migrants are represented by arrows as defined in the legend for 1990 (left) and 2008 (right). No migrants were inferred between management subunits or across rocky headlands. There is higher migration among the 2008 samples than 1990.

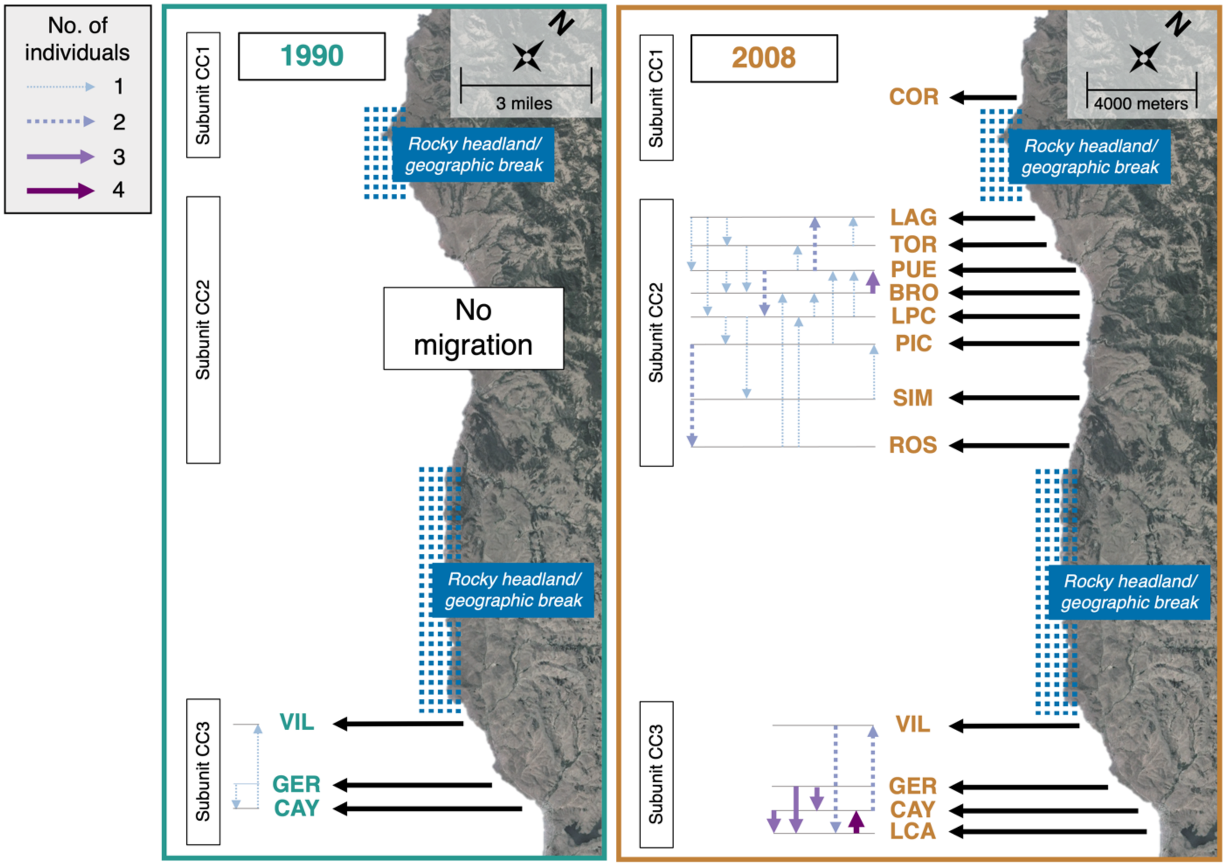


Figure 1.3: Expected heterozygosity (A) and average number of alleles per locus (B) for each respective site in 1990 (green) and 2008 (brown). Rocky headlands (blue dotted lines) separate management subunits. Three sites (BRO, ROS, and LCA) were only sampled in 2008. $H_{exp}$ for the 2008 Corral samples is significantly lower than the 1990 Corral samples (asterisk; paired Wilcoxon rank sum test, p-value < 0.005). Overall, there is a greater number of sites with higher $H_{exp}$ and an average number of alleles per locus in the 2008 samples compared to 1990. The 2008 samples are significantly higher in $H_{exp}$ compared to 1990 when including Corral (paired Wilcoxon rank sum test, p-value < 0.05) and it becomes more statistically significant when excluding Corral (paired Wilcoxon rank sum test, p-value < 0.0001). The 2008 samples also have a significantly higher average number of alleles per locus compared to the 1990 samples regardless of including or excluding Corral (paired Wilcoxon rank sum test; p-value < 0.0001).

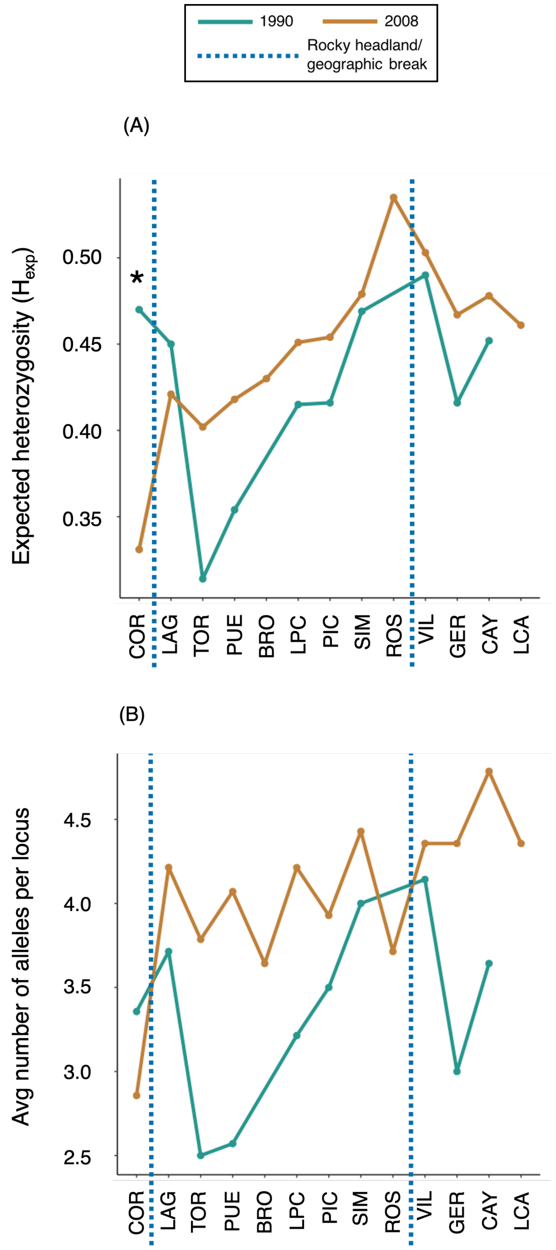


Figure 1.4: Total monthly precipitation in inches for the central coast for the 90 months preceding each sample collection month in 1990 (green) and 2008 (brown). Samples were collected in July and August 1990 and in February 2008. In 1990, there were 6 months with ≥4 inches of rainfall (bottom red dotted line) prior to the collection month whereas there were 11 such events prior to the 2008 collection month. One month prior to the 2008 sampling, there was 7 inches of rainfall (top red dotted line) whereas in 1990, there were 90 months prior to the 1990 sampling before such an event was recorded. Overall, there is a longer lapse in time between months with ≥4 inches of rainfall in 1990 compared to 2008. Rain gauge data was acquired from the Morro Bay Fire Department in California ([https://wrcc.dri.edu](https://wrcc.dri.edu)).

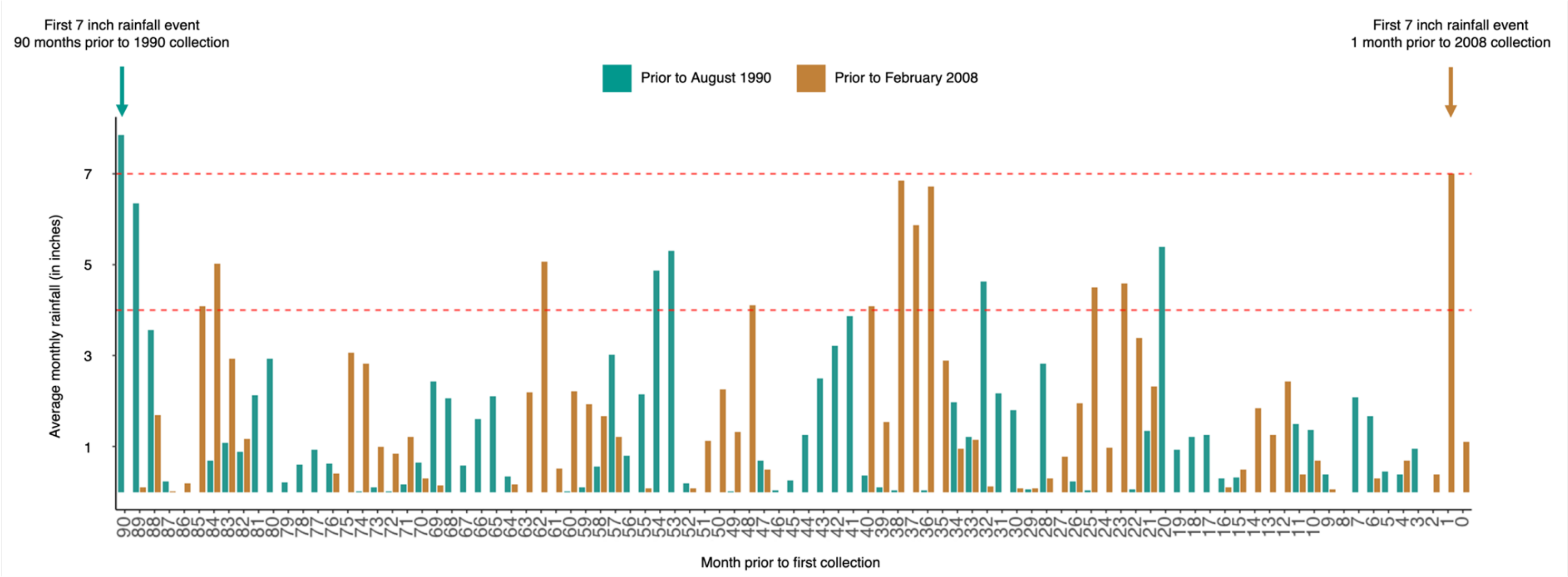
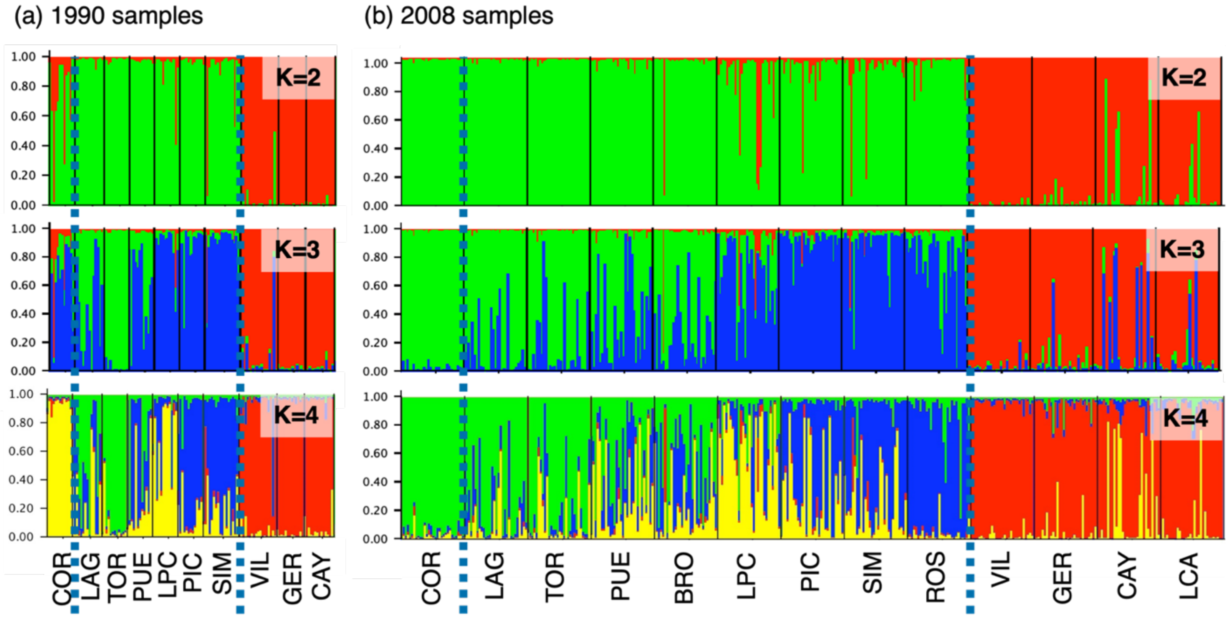


Figure 1.5: STRUCTURE plots at varying *K* values for each collection site in 1990 (a) and 2008 (b). Blue dotted lines delineate the two rocky headlands along the central coast that serve as large geographic breaks separating subunits from one another. Three sites (BRO, ROS, and LCA) were only sampled in 2008.



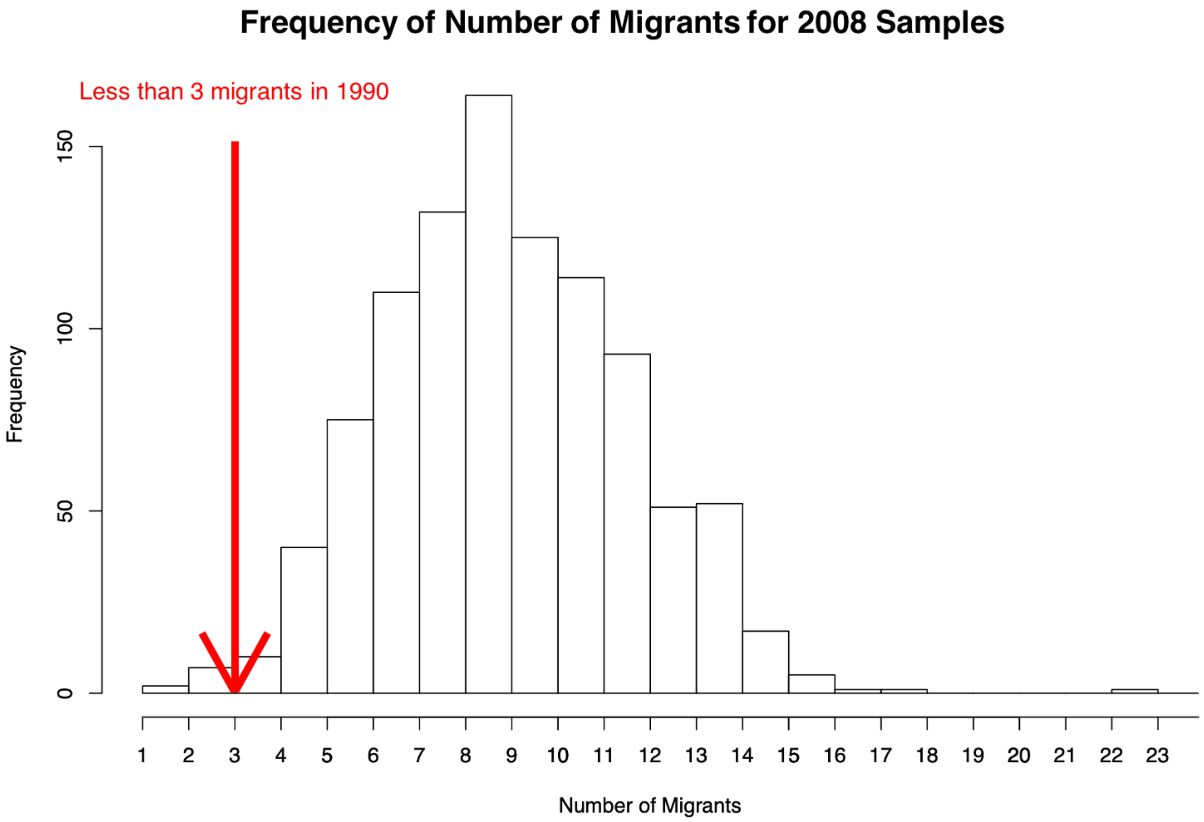
Appendix S1.1: Microsatellite loci for *Eucyclogobius*, including primer sequence (forward primer listed above reverse), length, cloned sequence repeat motif, size of range of alleles (base pairs), number of alleles observed.

| **Locus** | **Primer sequence (5' - 3')** | **Length** | **Motif** | **Fragment size (bp)** | **No. alleles** |
|---|---|---|---|---|---|
| ENE2 | GTCGACTGGCAGTATGGGAT | 20 | (ATCTCT)$_5$ | 142-154 | 3 |
| | AGACTCAAATATGTGCACACCAC | 13 | | | |
| ENE3 | CTAACCTGCCTTGACCCAAA | 20 | (GT)$_{12}$ | 120-130 | 2 |
| | GCAGAATCAACATAAGAACTATGGAAA | 27 | | | |
| ENE5 | GCTTGTGCAGTATGGGATCTC[a] | 21 | (GT)$_4$(AT)(GT)$_9$ | 306–326 | 5 |
| | CTCGGAGCGTTCATTTATCTC | 21 | | | |
| ENE6 | TCAGGTTTGTGCTAAAATGATG[a] | 22 | (CA)$_{11}$(CATACA)$_5$(CA)$_4$ | 241–257 | 10 |
| | TCCGATGACCACTTGTCC | 18 | | | |
| ENE7 | TCACATGAATCGGAGACAGT[a] | 20 | (CAT)$_7$(CAC)(CAT) | 135–159 | 6 |
| | CAGAGAGGGCACTTTTTCAG | 20 | | | |
| ENE8 | GAGGAAGGCGAGCTGATTA[a] | 19 | (ACCATCATC)4(ACC)(ATC)8 | 101–204 | 14 |
| | CGGAGAGAAGGTGTTGAGAG | 20 | | | |
| ENE9 | CCTTCATTTTTCCATCAGAAGCG[a] | 23 | (ATG)$_{28}$ | 131–209 | 27 |
| | CCTTATTTACATCTTCCCTCCA | 22 | | | |
| ENE12 | CTGGGATTGTCTTGGAACAG[a] | 20 | (GAT)$_9$ | 183–240 | 13 |
| | GGGTGTGTGTGAGAGAGTGG | 20 | | | |
| ENE15 | CCCGGAGGAGTTAGAGGAA[a] | 19 | (TGGA)$_7$ | 281–293 | 3 |
| | GAGCCTGTGGTTTGTCGAG | 19 | | | |
| ENE16 | GTCGCCTTGATTTTATTGTGA[a] | 21 | (TGGA)$_6$ | 138–216 | 6 |
| | CTCAGCGTGGTTTCATTAT | 19 | | | |
| ENE17 | CAGAGGTAGATCAGAAGAAC[a] | 20 | (ATCC)$_6$ | 165–173 | 2 |
| | CCGGATAAAGTGCAGAAAAT | 20 | | | |
| ENE18 | GGAGAACGAGAGAGAAAGA[a] | 19 | (GA)$_4$(AC)(GA)$_{25}$ | 132–152 | 8 |
| | GGCTGGTGTTTGATACATC | 19 | | | |
| ENE19 | CGCGTCAGTTTTCACCTTTA[a] | 20 | (TCTA)$_{11}$ | 110–134 | 3 |

| | | | | | |
|---|---|---|---|---|---|
| | GAGAATGCCCAAAATCACC | 19 | | | |
| ENE21 | TGCAGAGAAAGAGACAGGTATT[a] | 22 | $(\mathrm{CATCAA})_2(\mathrm{CAT})_7(\mathrm{CTG})_2(\mathrm{CAT})_3$ | 154–168 | 3 |
| | ATTGAGGTGCTGACACTGAG | 20 | | | |

[a] 23 nucleotide M13 sequence (5'-AGGGTTTTCCCAGTCACGACGTT-3') added to the 50 end of the forward primer to allow annealing of the dye-labeled M13 primer to the PCR product

Appendix S1.2: A rarefaction resampling assessment to confirm that the between year difference in assignment test results are not a function of the difference in sample sizes between sampling years. There were more samples collected in 2008 (n=390) compared to 1990 (n=136). For resampling, we generated 1000 subsets by randomly drawing microsatellite data from the larger 2008 data set equaling the number of individuals at each of the 10 sites in 1990 that were also sampled in 2008. This graph shows the distribution generated from the rarified data, which also provides a statistical distribution for assessing the significance of the observation of two migrants from the assignment tests (see Figure 1.2). Less than three migrants are observed in the rarified simulations seven times. Based on this simulation, the probability of this result occurring by chance is 7 out of 1000 (or p-value < 0.01). Thus, the difference in the number of estimated migrants across the two study years did not result from the variation in sample size.



Appendix S1.3: Data matrix with all pairwise $F_{ST}$ comparisons within the central coast management unit.

| | COR_08 | COR_90 | LAG_08 | LAG_90 | TOR_08 | TOR_90 | PUE_08 | PUE_90 | BRO_08 | LPC_08 | LPC_90 | PIC_08 | PIC_90 | SIM_08 | SIM_90 | ROS_08 | VIL_08 | VIL_90 | GER_08 | GER_90 | CAY_08 | CAY_90 | LCA_08 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| COR_08 | 0.000 | | | | | | | | | | | | | | | | | | | | | | | COR_08 |
| COR_90 | 0.122 | 0.000 | | | | | | | | | | | | | | | | | | | | | | COR_90 |
| LAG_08 | 0.040 | 0.087 | 0.000 | | | | | | | | | | | | | | | | | | | | | LAG_08 |
| LAG_90 | 0.060 | 0.077 | 0.026 | 0.000 | | | | | | | | | | | | | | | | | | | | LAG_90 |
| TOR_08 | 0.041 | 0.085 | 0.014 | 0.025 | 0.000 | | | | | | | | | | | | | | | | | | | TOR_08 |
| TOR_90 | 0.056 | 0.117 | 0.046 | 0.072 | 0.027 | 0.000 | | | | | | | | | | | | | | | | | | TOR_90 |
| PUE_08 | 0.054 | 0.074 | 0.024 | 0.030 | 0.027 | 0.046 | 0.000 | | | | | | | | | | | | | | | | | PUE_08 |
| PUE_90 | 0.144 | 0.100 | 0.090 | 0.084 | 0.091 | 0.116 | 0.051 | 0.000 | | | | | | | | | | | | | | | | PUE_90 |
| BRO_08 | 0.065 | 0.076 | 0.033 | 0.045 | 0.038 | 0.059 | 0.017 | 0.071 | 0.000 | | | | | | | | | | | | | | | BRO_08 |
| LPC_08 | 0.093 | 0.058 | 0.050 | 0.037 | 0.047 | 0.084 | 0.034 | 0.066 | 0.043 | 0.000 | | | | | | | | | | | | | | LPC_08 |
| LPC_90 | 0.110 | 0.063 | 0.063 | 0.052 | 0.055 | 0.095 | 0.041 | 0.079 | 0.052 | 0.015 | 0.000 | | | | | | | | | | | | | LPC_90 |
| PIC_08 | 0.109 | 0.063 | 0.058 | 0.033 | 0.048 | 0.093 | 0.041 | 0.074 | 0.055 | 0.017 | 0.028 | 0.000 | | | | | | | | | | | | PIC_08 |
| PIC_90 | 0.107 | 0.072 | 0.061 | 0.035 | 0.049 | 0.090 | 0.044 | 0.085 | 0.059 | 0.037 | 0.041 | 0.015 | 0.000 | | | | | | | | | | | PIC_90 |
| SIM_08 | 0.078 | 0.058 | 0.042 | 0.024 | 0.036 | 0.078 | 0.029 | 0.073 | 0.041 | 0.019 | 0.030 | 0.018 | 0.030 | 0.000 | | | | | | | | | | SIM_08 |
| SIM_90 | 0.084 | 0.061 | 0.046 | 0.030 | 0.039 | 0.073 | 0.034 | 0.071 | 0.044 | 0.021 | 0.036 | 0.021 | 0.032 | 0.011 | 0.000 | | | | | | | | | SIM_90 |
| ROS_08 | 0.124 | 0.087 | 0.090 | 0.082 | 0.094 | 0.142 | 0.096 | 0.123 | 0.094 | 0.069 | 0.093 | 0.085 | 0.108 | 0.062 | 0.060 | 0.000 | | | | | | | | ROS_08 |
| VIL_08 | 0.198 | 0.103 | 0.150 | 0.121 | 0.138 | 0.171 | 0.128 | 0.150 | 0.140 | 0.076 | 0.096 | 0.075 | 0.087 | 0.083 | 0.078 | 0.107 | 0.000 | | | | | | | VIL_08 |
| VIL_90 | 0.158 | 0.081 | 0.114 | 0.085 | 0.108 | 0.146 | 0.093 | 0.117 | 0.107 | 0.056 | 0.075 | 0.051 | 0.060 | 0.062 | 0.057 | 0.111 | 0.030 | 0.000 | | | | | | VIL_90 |
| GER_08 | 0.134 | 0.082 | 0.101 | 0.083 | 0.099 | 0.129 | 0.074 | 0.101 | 0.089 | 0.048 | 0.063 | 0.050 | 0.065 | 0.057 | 0.060 | 0.094 | 0.057 | 0.028 | 0.000 | | | | | GER_08 |
| GER_90 | 0.183 | 0.101 | 0.133 | 0.114 | 0.126 | 0.159 | 0.100 | 0.145 | 0.104 | 0.069 | 0.079 | 0.064 | 0.073 | 0.080 | 0.082 | 0.126 | 0.071 | 0.039 | 0.029 | 0.000 | | | | GER_90 |
| CAY_08 | 0.126 | 0.057 | 0.089 | 0.065 | 0.084 | 0.114 | 0.060 | 0.074 | 0.075 | 0.033 | 0.045 | 0.030 | 0.046 | 0.038 | 0.041 | 0.088 | 0.047 | 0.022 | 0.016 | 0.033 | 0.000 | | | CAY_08 |
| CAY_90 | 0.170 | 0.073 | 0.129 | 0.098 | 0.124 | 0.169 | 0.094 | 0.113 | 0.110 | 0.055 | 0.072 | 0.058 | 0.078 | 0.064 | 0.071 | 0.107 | 0.055 | 0.029 | 0.021 | 0.041 | 0.016 | 0.000 | | CAY_90 |
| LCA_08 | 0.140 | 0.074 | 0.102 | 0.079 | 0.101 | 0.143 | 0.075 | 0.096 | 0.084 | 0.042 | 0.060 | 0.042 | 0.056 | 0.049 | 0.053 | 0.093 | 0.051 | 0.024 | 0.012 | 0.025 | 0.013 | 0.015 | 0.000 | LCA_08 |

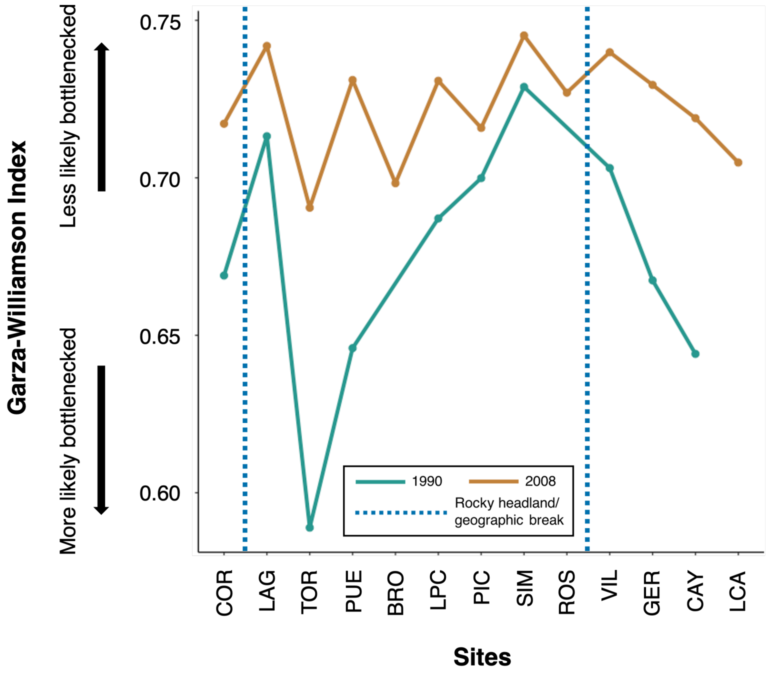
Appendix S1.4: Total monthly rainfall in inches 90 months prior (yellow) to the 1990 collection month (green) and the 2008 collection month (brown). The table shows data collected between 1980-1990 and 1998-2008 from a rain gauge at the Morro Bay Fire Department in California ([https://wrcc.dri.edu](https://wrcc.dri.edu)). Individual months are not used for annual statistics if more than five days are missing.

| **Year** | **Jan** | **Feb** | **Mar** | **Apr** | **May** | **Jun** | **Jul** | **Aug** | **Sep** | **Oct** | **Nov** | **Dec** | **Annual Total** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1980 | 6.35 | 5.83 | 2.32 | 0.8 | 0.36 | 0 | 0.4 | 0.02 | 0.05 | 0.07 | 0.01 | 0.5 | 16.71 |
| 1981 | 3.91 | 1.66 | 5.22 | 0.43 | 0.02 | 0 | 0 | 0.03 | 0 | 1.76 | 2.05 | 1.28 | 16.36 |
| 1982 | 2.62 | 2.19 | 6.53 | 3.26 | 0 | 0.29 | 0 | 0.01 | 1.02 | 1.17 | 4.39 | 1.5 | 22.98 |
| 1983 | 8.94 | 7.86 | 6.35 | 3.56 | 0.22 | 0 | 0 | 0.69 | 1.07 | 0.88 | 2.12 | 2.94 | 34.63 |
| 1984 | 0.21 | 0.6 | 0.92 | 0.63 | 0 | 0.01 | 0.1 | 0.02 | 0.16 | 0.64 | 2.43 | 2.05 | 7.77 |
| 1985 | 0.58 | 1.6 | 2.1 | 0.34 | 0 | 0 | 0 | 0.02 | 0.09 | 0.56 | 3.01 | 0.8 | 9.1 |
| 1986 | 2.15 | 4.86 | 5.31 | 0.18 | 0 | 0 | 0.01 | 0 | 0.68 | 0.03 | 0.26 | 1.25 | 14.73 |
| 1987 | 2.5 | 3.21 | 3.86 | 0.36 | 0.09 | 0.04 | 0 | 0.03 | 0 | 1.97 | 1.21 | 4.64 | 17.91 |
| 1988 | 2.17 | 1.79 | 0.06 | 2.83 | 0 | 0.22 | 0.03 | 0 | 0 | 0.06 | 1.34 | 5.4 | 13.9 |
| 1989 | 0.92 | 1.21 | 1.26 | 0.3 | 0.31 | 0 | 0 | 0 | 1.5 | 1.37 | 0.39 | 0 | 7.26 |
| 1990 | 2.08 | 1.67 | 0.44 | 0.38 | 0.95 | 0 | 0 | 0 | 0.44 | 0.02 | 0.28 | 0.57 | 6.83 |
| 1998 | 4.62 | 11.28 | 4.6 | 1.72 | 2.46 | 0 | 0 | 0 | 0.12 | 0.49 | 1.65 | 0.56 | 27.5 |
| 1999 | 2.51 | 1.86 | 3.92 | 1.96 | 0 | 0 | 0 | 0 | 0.74 | 0.07 | 1.29 | 0.09 | 12.44 |
| 2000 | 3.65 | 9.92 | 1.32 | 2.23 | 0.28 | 0.38 | 0 | 0 | 0.1 | 1.68 | 0.02 | 0.19 | 19.77 |
| 2001 | 4.08 | 5.03 | 2.92 | 1.17 | 0 | 0 | 0 | 0 | 0 | 0.4 | 3.07 | 2.82 | 19.49 |
| 2002 | 0.99 | 0.84 | 1.2 | 0.3 | 0.14 | 0 | 0 | 0 | 0 | 0.17 | 2.18 | 5.07 | 10.89 |
| 2003 | 0.52 | 2.22 | 1.92 | 1.67 | 1.2 | 0 | 0.08 | 0 | 0 | 0.08 | 1.13 | 2.26 | 11.08 |
| 2004 | 1.32 | 4.11 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 4.09 | 1.54 | 6.86 | 18.42 |
| 2005 | 5.88 | 6.73 | 2.89 | 0.94 | 1.14 | 0.12 | 0 | 0.08 | 0.08 | 0.29 | 0.77 | 1.94 | 20.86 |
| 2006 | 4.5 | 0.97 | 4.58 | 3.39 | 2.33 | 0 | 0 | 0 | 0 | 0.1 | 0.49 | 1.84 | 18.2 |
| 2007 | 1.26 | 2.42 | 0.38 | 0.69 | 0.06 | 0 | 0 | 0.29 | 0 | 0.69 | 0 | 0.39 | 6.18 |
| 2008 | 7 | 1.1 | 0.03 | 0.56 | 0 | 0 | 0 | 0.03 | 0 | 0.2 | 1 | 1.58 | 10.17 |

Appendix S1.5: Z-score for within-site interannual $F_{ST}$ comparison. Corral is significantly higher than all other sites (p-value = 0.007).

| Site | Year | Management_unit | Paired_Site_Name | Pairwise_Fst | Fst_zscore |
|---|---|---|---|---:|---:|
| Corral | 1990 and 2008 | Central Coast Unit (CCU) | Corral | 0.122 | 2.66597005 |
| Laguna | 1990 and 2008 | Central Coast Unit (CCU) | Laguna | 0.026 | -0.2489858 |
| Tortuga | 1990 and 2008 | Central Coast Unit (CCU) | Tortuga | 0.027 | -0.2186217 |
| Puerto | 1990 and 2008 | Central Coast Unit (CCU) | Puerto | 0.051 | 0.51011728 |
| Little Pico | 1990 and 2008 | Central Coast Unit (CCU) | Little Pico | 0.015 | -0.5829912 |
| Pico | 1990 and 2008 | Central Coast Unit (CCU) | Pico | 0.015 | -0.5829912 |
| San Simeon | 1990 and 2008 | Central Coast Unit (CCU) | San Simeon | 0.011 | -0.7044477 |
| Villa | 1990 and 2008 | Central Coast Unit (CCU) | Villa | 0.030 | -0.1275293 |
| San Geronimo | 1990 and 2008 | Central Coast Unit (CCU) | San Geronimo | 0.029 | -0.1578934 |
| Cayucos | 1990 and 2008 | Central Coast Unit (CCU) | Cayucos | 0.016 | -0.552627 |
| | | | | | |
| Mean | 0.034 | | | | |
| St Dev | 0.032933603 | | | | |

Appendix S1.6: Garza-Williamson (G-W; Garza and Williamson 2001) index was calculated for all sites in 1990 (green) and 2008 (brown). Three sites (BRO, ROS, and LCA) were only sampled in 2008. The G-W index is based on the ratio of the number of microsatellite alleles to the range in allele size to detect genetic bottlenecks. Two rocky headlands (blue dotted lines) serve as large geographic barriers between subunits within the Central Coast management unit (CCU). The G-W index is significantly higher in 2008 than 1990 for the entire dataset (t-test; p-value < 0.0005); however, we suspect the G-W index is less effective than other measures of genetic variation in identifying bottlenecks in this context due to extensive subdivision and complex history of tidewater gobies.



Appendix S1.7: Arroyo del Corral, San Luis Obispo County. This photo illustrates the severe impacts of drought and desiccation on small lagoons, which are critical habitats for tidewater gobies. Photo taken by Brenton T. Spies on July 9, 2015.

## References

Behrens, D. K., Bombardelli, F. A., Largier, J. L., & Twohy, E. (2013). Episodic closure of the tidal inlet at the mouth of the Russian River—A small bar-built estuary in California. *Geomorphology*, *189*, 66–80. https://doi.org/10.1016/j.geomorph.2013.01.017

Berg, N., & Hall, A. (2015). Increased interannual precipitation extremes over California under climate change. *Journal of Climate*, *28*(16), 6324–6334.

Bond, M. H., Hayes, S. A., Hanson, C. V., & MacFarlane, R. B. (2008). Marine survival of steelhead (Oncorhynchus mykiss) enhanced by a seasonally closed estuary. *Canadian Journal of Fisheries and Aquatic Sciences*, *65*(10), 2242–2252. https://doi.org/10.1139/F08-131

Boutin-Ganache, I., Raposo, M., Raymond, M., & Deschepper, C. F. (2001). M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *Biotechniques*, *31*(1), 25–28.

Dawson, M. N., Louie, K. D., Barlow, M., Jacobs, D. K., & Swift, C. C. (2002). Comparative phylogeography of sympatric sister species, Clevelandia ios and Eucyclogobius newberryi (Teleostei, Gobiidae), across the California Transition Zone. *Molecular Ecology*, *11*(6), 1065–1075.

Dawson, M. N., Staton, J. L., & Jacobs, D. K. (2001). Phylogeography of the Tidewater Goby, Eucyclogobius newberryi (Teleostei, Gobiidae), in Coastal California. *Evolution*, *55*(6), 1167–1179.

Earl, D. A., Louie, K. D., Bardeleben, C., Swift, C. C., & Jacobs, D. K. (2010). Rangewide microsatellite phylogeography of the endangered tidewater goby, Eucyclogobius newberryi (Teleostei: Gobiidae), a genetically subdivided coastal fish with limited marine dispersal. *Conservation Genetics*, *11*(1), 103–114. https://doi.org/10.1007/s10592-009-0008-9

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, *1*, 117693430500100. https://doi.org/10.1177/117693430500100003

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567.

García-Reyes, M., & Largier, J. L. (2012). Seasonality of coastal upwelling off central and northern California: New insights, including temporal and spatial variability. *Journal of Geophysical Research: Oceans*, *117*(C3).

Garza, J. C., & Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, *10*(2), 305–318. https://doi.org/10.1046/j.1365-294x.2001.01190.x

Hayes, S. A., Bond, M. H., Hanson, C. V., Freund, E. V., Smith, J. J., Anderson, E. C., Ammann, A. J., & MacFarlane, R. B. (2008). Steelhead Growth in a Small Central California Watershed: Upstream and Estuarine Rearing Patterns. *Transactions of the American Fisheries Society*, *137*(1), 114–128. https://doi.org/10.1577/T07-043.1

Hoeksema, S. D., Chuwen, B. M., Tweedley, J. R., & Potter, I. C. (2018). Factors influencing marked variations in the frequency and timing of bar breaching and salinity and oxygen regimes among normally-closed estuaries. *Estuarine, Coastal and Shelf Science*, *208*, 205–218. https://doi.org/10.1016/j.ecss.2018.04.010

Huber, E. R., & Carlson, S. M. (2020). Environmental correlates of fine-scale juvenile steelhead trout (Oncorhynchus mykiss) habitat use and movement patterns in an intermittent estuary during drought. *Environmental Biology of Fishes*, *103*(5), 509–529. https://doi.org/10.1007/s10641-020-00971-y

Huyer, A., Smith, R. L., & Fleischbein, J. (2002). The coastal ocean off Oregon and northern California during the 1997–8 El Nino. *Progress in Oceanography*, *54*(1–4), 311–341.

Jacobs, D., Stein, E. D., & Longcore, T. (2011). Classification of California estuaries based on natural closure patterns: Templates for restoration and management. *Southern California Coastal Water Research Project*, *619*, 1–50.

Kinziger, A. P., Hellmair, M., McCraney, W. T., Jacobs, D. K., & Goldsmith, G. (2015). Temporal genetic analysis of the endangered tidewater goby: Extinction–colonization dynamics or drift in isolation? *Molecular Ecology*, *24*(22), 5544–5560.

Kraus, N. C. (2003). Analytical model of incipient breaching of coastal barriers. *Coastal Engineering Journal*, *45*(04), 511–531.

Lafferty, K. D., & Page, C. J. (1997). Predation on the endangered tidewater goby, Eucyclogobius newberryi, by the introduced African clawed frog, Xenopus laevis, with notes on the frog's parasites. *Copeia*, *1997*(3), 589–592.

Lafferty, K. D., Swenson, R. O., & Swift, C. C. (1996). Threatened fishes of the world: Eucyclogobius newberryi Girard, 1857 (Gobiidae). *Environmental Biology of Fishes*, *46*(3), 254–254. https://doi.org/10.1007/BF00005000

Lafferty, K. D., Swift, C. C., & Ambrose, R. F. (1999a). Extirpation and Recolonization in a Metapopulation of an Endangered Fish, the Tidewater Goby. *Conservation Biology*, *13*(6), 1447–1453. https://doi.org/10.1046/j.1523-1739.1999.98016.x

Lafferty, K. D., Swift, C. C., & Ambrose, R. F. (1999b). Postflood persistence and recolonization of endangered tidewater goby populations. *North American Journal of Fisheries Management*, *19*(2), 618–622.

McCraney, W. T., Goldsmith, G., Jacobs, D. K., & Kinziger, A. P. (2010). Rampant drift in artificially fragmented populations of the endangered tidewater goby (Eucyclogobius newberryi). *Molecular Ecology*, *19*(16), 3315–3327.

Mendonca, H., Smith, J., & Brinegar, C. (2001). Isolation and characterization of four microsatellite loci in the tidewater goby (Eucyclogobius newberryi). *Marine Biotechnology*, *3*(2), 91–95.

National Weather Service Climate Prediction Center, College Park. Available from https://cpc.ncep.noaa.gov

Orescanin, M. M., & Schooler, J. (2018). Observations of episodic breaching and closure at an ephemeral river. *Continental Shelf Research*, *166*, 77–82.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pritchard, J. K., Wen, W., & Falush, D. (2010). Documentation for STRUCTURE software: Version 2. *University of Chicago, Chicago, IL*.

R Core Team, R. (2013). *R: A language and environment for statistical computing*.

Ranasinghe, R., & Pattiaratchi, C. (2003). The seasonal closure of tidal inlets: Causes and effects. *Coastal Engineering Journal*, *45*(4), 601–627.

Reese, D. A., & Welsh Jr, H. H. (1998). Habitat use by western pond turtles in the Trinity River, California. *The Journal of Wildlife Management*, 842–853.

Reis, D. K. (1999). *Habitat characteristics of California red-legged frogs (Rana aurora draytonii): Ecological differences between eggs tadpoles, and adults in a coastal brackish and freshwater system*. 70.

Rich, A., & Keller, E. A. (2013). A hydrologic and geomorphic model of estuary breaching and closure. *Geomorphology*, *191*, 64–74. https://doi.org/10.1016/j.geomorph.2013.03.003

Schooler, J. D. (2017). *Episodic changes in lagoon water levels due to ephemeral river breaching and closure events*. Naval Postgraduate School Monterey United States.

Spies, B. T., & Steele, M. A. (2016). Effects of temperature and latitude on larval traits of two estuarine fishes in differing estuary types. *Marine Ecology Progress Series*, *544*, 243–255.

Swenson, R. O. (1997). Sex-role reversal in the tidewater goby, Eucyclogobius newberryi. *Environmental Biology of Fishes*, *50*, 27–40.

Swenson, R. O. (1999). The ecology, behavior, and conservation of the tidewater goby, Eucyclogobius newberryi. *Environmental Biology of Fishes*, *55*(1), 99–114.

Swift, C. C., Haglund, T. R., Ruiz, M., & Fisher, R. N. (1993). The status and distribution of the freshwater fishes of southern California. *Bulletin of the Southern California Academy of Sciences*, *92*(3), 101–167.

Swift, C. C., Mulder, J., Dellith, C., & Kittleson, K. (2018). Mortality of Native and Non-native Fishes during Artificial Breaching of Coastal Lagoons in Southern and Central California. *Bulletin, Southern California Academy of Sciences*, *117*(3), 157–168. https://doi.org/10.3160/1767.1

Swift, C. C., Nelson, J. L., Maslow, C., & Stein, T. (1989). *Biology and Distribution of the Tidewater Goby, Eucyclogobius Newberryi (Pisces: Gobiidae) of California*.

Swift, C. C., Spies, B., Ellingson, R. A., & Jacobs, D. K. (2016). A new species of the bay goby genus Eucyclogobius, endemic to southern California: Evolution, conservation, and decline. *PloS One*, *11*(7), e0158543.

Thuiller, W. (2007). Climate change and the ecologist. *Nature*, *448*(7153), 550–552.

U.S. Fish and Wildlife Services. (2005). *Recovery plan for the tidewater goby (Eucyclogobius newberryi)*. https://www.fws.gov/pacific/ecoservices/endangered/recovery/documents/TidewaterGobyfinalRecoveryPlan.pdf

U.S. Fish and Wildlife Services. (2013). *Endangered and Threatened Wildlife and Plants; Designation of Critical Habitat for Tidewater Goby*. https://www.govinfo.gov/content/pkg/FR-2013-02-06/pdf/2013-02057.pdf

U.S. Fish and Wildlife Services. (2014). *Endangered and Threatened Wildlife and Plants; Reclassifying the Tidewater Goby from Endangered to Threatened*. https://www.govinfo.gov/content/pkg/FR-2014-03-13/pdf/2014-05335.pdf

Western Regional Climate Center, Reno. Available from https://wrcc.dri.edu

Williams, M. (2014). *Hydrodynamics and Salt Dispersion in Intermittently Closed Bar-Built Estuaries* [UC Berkeley]. https://escholarship.org/uc/item/4h3770rr

# Chapter 2

# Implications for how landscape genetics inform metapopulation dynamics and conservation for the endangered tidewater goby (Genus: *Eucyclogobius*)

Benjamin A. Hà,[1] Jonathan Y. Sim,[1] Dan Holland,[1] Camm C. Swift,[2]
Kyle Cavanaugh,[3] David K. Jacobs[1]

[1] Department of Ecology and Evolutionary Biology, University of California, Los Angeles
[2] Natural History Museum of Los Angeles County
[3] Department of Geography, University of California, Los Angeles

## Abstract

Landscape genetic studies have mostly focused on terrestrial systems. Seascape genetics has been less studied by comparison. There also seems to be a general need to explore how landscape genetics inform metapopulation dynamics. The endangered tidewater gobies (Genus: *Eucyclogobius*) are exclusive to California coastal lagoons where they exhibit metapopulation dynamics that depend on stream hydrology. Dispersal occurs between lagoons and is influenced by coastal properties. Tidewater gobies represent a promising model to understand how landscape/seascape genetics can inform conservation management and metapopulation dynamics, such as dispersal and source-sink processes. Here, we analyzed data from 14 microsatellite loci for 1346 tidewater goby individuals covering four management units over an 18-year timespan. Data for landscape variables include geographic distance, the proportion of seven soft substrate and hard rock types, and the number of headlands. Seascape variables include the average area of kelp vegetation. All variables were collected along the California coast by analyzing remote sensing imagery and geologic maps. Results from multiple linear regression (MLR) models provide evidence that heterogeneous landscapes that include softer substrates facilitate dispersal between sites. There is also evidence that harder substrates, the presence of headlands, and kelp vegetation may reduce gene flow. Based on population genetic statistics and variability in population structures, results suggest the subunits within the Conception Management Unit (COU) slightly differ from the current subunit designations defined by U.S. Fish and Wildlife Service. We provide an assessment of how these landscape/seascape variables relate to the genetic structures and dispersal, and consequently the metapopulation dynamics, between tidewater goby populations. We conclude by providing recommendations for defining management subunits for *Eucyclogobius*.

## Introduction

Landscape genetics research can enhance our understanding of the interaction between landscape variables and evolutionary processes, such as gene flow, genetic drift, and adaptation (Manel et al., 2003; Manel & Holderegger, 2013). Understanding the effects of landscape variables on genetic connectivity between populations can also shed light on fundamental biological processes, such as metapopulation dynamics, speciation, dispersal, and population structures (Jenkins et al., 2010; Rissler, 2016; Storfer et al., 2007). This area of research becomes complex given that populations inhabit heterogeneous landscapes, where environmental factors can promote or impede the dispersal of individuals between populations. In turn, landscape genetic studies can have important applications for conservation management. For example, identifying landscape features, especially when caused anthropogenically, that reduce habitat connectivity and increase habitat fragmentation can promote conservation action for mitigating negative impacts (Blair & Melnick, 2012; Kunde et al., 2020; Segelbacher et al., 2010; Storfer et al., 2007).

Since the term "landscape genetics" was first coined in 2003 (Manel et al., 2003), the number of relevant studies has continued to increase, especially given the rapid advances in our ability to acquire molecular genetic data and high-resolution landscape data (Storfer et al., 2010). Several studies have highlighted the effects of landscape genetics across different systems, including mammals (Blair & Melnick, 2012; Flores-Manzanero et al., 2019; Kunde et al., 2020; Mapelli et al., 2020; Razgour et al., 2014; Waits et al., 2015), birds (Bicknell et al., 2012; Leon et al., 2022; Morinha et al., 2017; Ramos et al., 2016), plants (Holderegger et al., 2010; Sork & Smouse, 2006), and amphibians (Emaresi et al., 2011; Haugen et al., 2020; Homola et al., 2019).

Although landscape genetics research has focused more on terrestrial habitats with less attention on landscape or seascape genetics research on aquatic or coastal systems (Storfer et al., 2010). There have been relevant studies across diverse aquatic environments (Riginos & Liggins, 2013; Selkoe et al., 2015) including a rising interest in seascape genetics research on marine organisms, including fish, dolphins, and invertebrates (Amaral et al., 2012; Cook et al., 2011; Leclerc et al., 2008; López-Márquez et al., 2021; Selkoe et al., 2008, 2016). Additionally, few landscape genetics studies have addressed specific hypotheses related to connectivity, such as identifying source-sink processes (Storfer et al., 2010).

In this respect, tidewater gobies serve as a promising model to expand landscape and seascape genetics research (hereafter, used interchangeably as landscape genetics). The federally endangered tidewater gobies (Genus: *Eucyclogobius*) exhibit metapopulation dynamics that are dependent on coastal hydrologic processes (U.S. Fish and Wildlife Services, 1994, 2005, 2013, 2014). They live exclusively in shallow estuaries and lagoons along the California coast, where movement is strictly dependent on heavy rainfall events and the subsequent breaching of lagoons, resulting in intermittent dispersal. Source-sink populations have been identified mostly through fieldwork (Lafferty et al., 1999a; U.S. Fish and Wildlife Services, 2005) in addition to genetic evidence for extinction-recolonization processes (Martel et al., 2021) (see Chapter 1 "Hydrologic Control of Metapopulation Dispersal in the Endangered Northern Tidewater Goby (*Eucyclogobius newberryi*) on the California Coast"). The intermittent dispersal of tidewater gobies, in addition to the seasonal dynamics of the habitat, contributes to high local genetic differentiation within and between metapopulations, creating a hierarchy of genetic subdivisions (Dawson et al., 2001a, 2002; Earl et al., 2010; Kinziger et al., 2015; McCraney et al., 2010). This has helped inform the designations of management units and subunits for the genus (U.S. Fish and Wildlife Services, 2005, 2014), where dispersal is assumed to be very limited temporally and largely restricted spatially to within subunits.

During dry periods, such as in most summer months, lagoons are closed to the sea, and tidewater gobies are geographically and genetically isolated. In extreme cases, desiccation can lead to habitat elimination and population extirpation (Lafferty et al., 1999a; Swift et al., 1993). Other factors can also alter population sizes and habitats, such as invasive species, anthropogenic development, and artificial filling and opening of habitats (Lafferty et al., 1996, 1999a; U.S. Fish and Wildlife Services, 2005, 2014). Reproductive periods for tidewater gobies occur during the summer (i.e., no to little rainfall), which suggests tidewater gobies likely disperse as adults since they lack a marine larval life history stage (Lafferty et al., 1999b, 1999c; Spies, 2014; Swenson, 1999). This has led to the hypothesis that tidewater gobies have low dispersal rates (Barlow, Michele, 2002; Dawson et al., 2001a; Earl et al., 2010) since dispersal is restricted to rainfall during the winter months.

Potentially exploring landscape genetics studies for tidewater gobies is simplified by the linear arrangement of coastal habitats in California. The closest studies related to landscape genetics and tidewater gobies include phylogeographic research using mitochondrial DNA and microsatellites to assess the relation between phylogenetics and biogeography. Evidence suggests tidewater gobies are absent where precipitous coastlines preclude suitable habitat from forming (Dawson et al., 2001b), which also forms phylogeographic structures as a result of life history strategies (Dawson et al., 2002; Earl et al., 2010; Swift et al., 1989). Similar phylogenetic breaks have been found based on morphology (Ahnelt et al., 2004). Geographic headlands, hard rocky substrates, and long stretches in distance between habitats may limit dispersal, forming local and regional phylogeographic breaks (Barlow, Michele, 2002; Dawson et al., 2001a; Earl et al., 2010). This has suggested that harder substrates and headlands restrict gene flow compared to sites separated by continuous, softer substrates (Barlow, Michele, 2002; Dawson et al., 2001a; U.S. Fish and Wildlife Services, 2005). However, there have yet to be studies that assess how landscape variables relate to metapopulation dynamics and affect population structures spatially within and between management units. Here, we investigate the following questions: 1) what is the relation between landscape variables and population genetic distance metrics, 2) how do landscape variables relate to tidewater goby metapopulation dynamics, such as dispersal and source-sink processes, and 3) what are the conservation implications for *Eucyclogobius*?

## Methods

### *Sample sites and collections*

A total of 1346 tidewater goby individuals were collected by seine net (4.5x2 meter with 20 cm mesh) from 40 total unique sites covering four management units across an 18-year timespan (Table 2.1). Samples from 23 sites were only collected in one year (1999, 2001, 2002, 2006, or 2008) and collectively represent two management units: Conception Unit (COU) and Los Angeles/Ventura Unit (LAVU). The remaining 17 sites have samples mostly collected from two separate years and collectively represent two additional management units: Central Coast Unit (CCU) and South Coast Unit (SCU). Samples from 10 sites (Arroyo del Corral, COR; Arroyo Laguna, LAG; Arroyo de Tortuga, TOR; Arroyo del Puerto, PUE; Little Pico Creek, LPC; Pico Creek, PIC; San Simeon Creek, SIM; Villa Creek, VIL; San Geronimo Creek, GER; Cayucos Creek, CAY) were collected in 1990 and 2008 in the CCU, and samples from two sites (San Onofre Creek, ONO; Las Flores Creek, FLO) were collected in 1990 and 2007 in the SCU. The exceptions are three sites in CCU (Broken Bridge Creek, BRO; Santa Rosa Creek, ROS;

Little Cayucos Creek, LCA), which were only sampled in 2008, and two sites in SCU (Aliso Canyon, ALI; Santa Margarita River, MARG), which were only sampled in 2007 and 1990, respectfully.

Samples were collected by D. Holland in 1990 prior to the federal listing of tidewater gobies under the Endangered Species Act with permission from the California Department of Fish and Game. All other samples were collected by D.K. Jacobs and C.C. Swift under a federal permit to C.C. Swift from the California Department of Fish and Game (Permit # SCP-2679) and U.S. Fish and Wildlife Service (Permit #TE-793644).

### *DNA extraction and genotyping*

Tidewater gobies were preserved on ice in 95% ethanol or liquid nitrogen in the field before being stored at -80°C until DNA extraction. Whole genomic DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit following the manufacturer's instructions. Tidewater gobies were genotyped at 14 total microsatellite loci (Appendix S2.1, Supporting Information) (Earl et al., 2010; Mendonca et al., 2001). Microsatellite primers were based on Mendonca *et al.* 2001 and Earl *et al.* 2010, and thermal cycling conditions and reaction volumes are described in Earl *et al.* 2010. Products were run on an ABI 3730KL capillary sequencer and scored using ABI GENEMAPPER version 3.0 (Applied Biosystems).

### *Data Analysis: Genetic distance metrics and genetic structures between populations*

GenAlEx 6.5 (Peakall & Smouse, 2012) was used to calculate expected heterozygosity ($H_{exp}$), the average number of alleles per locus, pairwise $F_{ST}$, Nei's genetic distance, and Nei's unbiased genetic distance, where the latter three items are all different metrics for genetic distance between pairs of populations. Genetic distance and $F_{ST}$ (Wright, 1951) are often used to measure genetic variation between populations. Since $F_{ST}$ can become inflated by factors that reduce within-population genetic variation (Charlesworth, 1998; Nei, 1973), $F_{ST}$ can function more like a measure of inbreeding within subsamples than a genetic distance between subsample pairs (Sere et al., 2017). Therefore, Nei's genetic distance and Nei's unbiased genetic distance (Nei, 1973, 1978) were used to measure genetic differentiation between population pairs in the study; however, all genetic distance metrics are reported. Additionally, the sample sizes for our study align with the suggestion that about 20 individuals per population should be sufficient when $F_{ST}$ was greater than 0.05 (Kalinowski, 2005).

STRUCTURE 2.3 (Pritchard et al., 2010) was used to estimate the number of population subdivisions (*K*) using Bayesian clustering from microsatellite data. We used STRUCTURE in the following ways: on the full data set and within each of the four management units. To address the same unique sites collected in two separate years within CCU and SCU, respectively, we assigned the same and different putative population origin for each individual. For example, the samples from Arroyo del Corral collected in 1990 and 2008 in the CCU were assigned to the same putative population (both 1990 and 2008 as population #1) and different putative populations (1990 as population #1 and 2008 as population #2). We employed the admixture model with 20 replicates for each *K* number of clusters permitted in the analysis between 2 and 40 using a burn-in of 100,000 iterations followed by 100,000 Markov-Chain Monte-Carlo (MCMC) steps. STRUCTURE HARVESTER (http://taylor0.biology.ucla.edu/struct_harvest) was used to assess and visualize the likelihood values across multiple values of *K* to identify the number of genetic groups that best fit the data estimated by the optimum *K* value (Earl & vonHoldt, 2012). Fourteen total microsatellite loci were used across all samples; however, there was some variation in loci with missing data across all management units (Table 2.1). Microsatellite loci were not included in the analysis if ≥50% of the loci had missing data.

The *adegenet* package (Jombart, 2008) in R was used to assess the discriminant analysis of principal components (DAPC). DAPC is a multivariate method similar to a principal components analysis (PCA) that identifies clusters of genetically related individuals. DAPC also produces linear discriminant (LD) outputs (e.g., LD1, LD2) that explain a proportion of the variance calculated from eigenvalues similar to a principal component (PC) output (e.g., PC1, PC2). However, unlike a PCA, a DAPC allows for a group assessment using a *K*-means value to infer a specific number of genetic clusters. Additionally, a DAPC focuses on between-group variability while neglecting within-group variation (Jombart et al., 2010), which may be more suitable for assessing population genetic structures for this study. DAPC analysis has also been demonstrated to perform better than STRUCTURE in characterizing population subdivisions (Jombart et al., 2010). When performing a DAPC for loci with missing data, the missing data are replaced by the mean allele frequency (Jombart, 2008; Jombart & Collins, n.d.).

### *Data collection for landscape variables*

All landscape variables were determined for each pair of lagoons where microsatellite genetic observations within management units had been generated along the coast. For analysis between management units, landscape variable data included the pairwise sites for those residing at the edge connecting any two adjacent management units (e.g., Little Cayucos from the CCU to San Luis Obispo Creek in the COU) (Table 2.1). The following landscape variables were assessed: total geographic distance, the number of headlands, the proportion of seven different types of soft and hard substrates, and the average area of kelp. Further description for each landscape variable is described below.

When collecting data on the landscape variables that required distance measurements, specifically geographic distance and the proportion of different soft and hard substrates, we had to account for the coastline paradox. This is a mathematical concept that describes a counterintuitive observation that the length of a coastline is not a fixed quantity, but rather depends on the scale at which it is measured (i.e., a coastline appears to be much longer when measured by mm than it does when measured by km). To account for this, we defined a protocol to collect distance measurements using relatively consistent unit increments between adjacent pairwise sites.

The total geographic distance and the number of rocky headlands between pairwise sites were collected via Google Earth Pro (Landsat 7). Using the ruler feature in Google Earth Pro, assessments for geographic distance between sites were measured in increments ranging from 300 (min) to 500 (max) meters; the total geographic distance between all pairwise sites was the sum of these increments. These measurements were taken parallel to the coast about 250 to 350 meters away from the coastline. The number of headlands was also quantified. A headland was identified via Google Earth Pro (Landsat 7) based on visual estimation that conformed to well-known coastal features (Appendix S2.2, Supporting Information).

Soft substrate and hard rock type exposure in the intertidal/adjacent subtidal zones were assessed using a coastal time series in Google Earth Pro, aerial imagery from the California Coastal Records Project database (https://www.californiacoastline.org), and Dibblee Geologic Maps (https://www.sbnaturestore.org; Cambria: DF-364; Carpinteria: DF-04; Casmalia: DF-24; Cayucos: DF-216; Dos Pueblos: DF-09; Gaviota: DF-16; Goleta: DF-07; Morro Bay North: DF-

215; Morro Bay South: DF-214; Piedras Blancas: DF-367; Pismo Beach: DF-212; Point Conception: DF-18; Port San Luis: DF-213; Point Arguello: DF-19; Point Sal: DF-25; Sacate: DF-17; San Simeon: DF-366; Santa Barbara: DF-06; Surf: DF-20; Tajiguas: DF-15). From the time series available in the image sources, those with good intertidal exposure were used (i.e., low tide) and the series was roughly assessed for seasonal variation in sand cover. The following seven categories were delineated from softest to hardest substrate: 1) Perennial sand (SP), 2) sand spatially and/or temporarily mixed with rock (SR), 3) erosion of terrace deposits (TD), 4) general exposure of soft, fine sedimentary bedded rocks (FG), 5) hard sandstones (HC), 6) extrusive volcanic rocks (VO), and 7) intrusive volcanic rocks (IN). Using the ruler feature, assessments for each of the seven substrates were measured in increments ranging from 300 (min) to 500 (max) meters. Each coastline point had been assigned a substrate type as defined by the Dibblee Geologic Maps. These were then used to determine the proportional distance of each amount of substrate type present relative to one another for each pairwise site. The final estimated distance of each substrate type was calculated based on the total geographic distance between pairwise sites collected previously.

KelpWatch (https://kelp.codefornature.org) was used to quantify the average area of kelp canopy ($m^2$) between sites by using a time series derived from five satellite Landsat sensors (Bell et al., 2023). Kelp area was identified by drawing a polygon that spanned the region between pairwise sites and extracting the area of kelp canopy within this polygon. The measurements were downloaded as a .csv file to calculate the average area of kelp between 1990 and 2008, the timespan for all samples collected across all sites.

### *Data Analysis: Landscape variables*

JMP software (version 17.0.0) was used to run a multiple linear regression (MLR) model to identify how the predictor variables (i.e., landscape variables) relate to a response variable (i.e., a population genetic distance metric). JMP was also used to calculate Akaike information criterion (AIC) to compare different MLRs with the same data input. Outputs for effect sizes of landscape variables include eta squared, partial eta squared, and omega squared. Eta squared ($\eta^2$) ranges from 0 to 1 and measures the proportion of the total variance in a dependent variable that is associated with different groups defined by an independent variable in a sample ($\eta^2 < 0.01$ considered small effect size, $0.01 < \eta^2 < 0.06$ considered medium, and > 0.14 considered large) (Richardson, 2011). Partial eta squared ($\eta^2_p$) is a similar measure in which the effects of other independent variables and interactions of variables are partially omitted. Omega squared ($\omega^2$) ranges from 0 to 1 and is an estimate of the amount of variance in the dependent variable that is accounted for by the independent variables in a population ($\omega^2 < 0.01$ considered very small, $0.01 < \omega^2 < 0.06$ considered small, $0.06 < \omega^2 < 0.14$ considered medium, > 0.14 considered large). All three metrics for effect size are reported.

### *Additional statistical analyses*

Graph generation and general statistical analyses (Pearson's correlation coefficient, coefficient of determination, z-scores) were performed in R (R Core Team, 2013). Pearson's correlation coefficient ($r$) is a value that ranges from -1 to 1 and measures the strength and directionality of a correlation between two variables ($r < 0.29$ considered small, $0.3 < r < 0.49$ considered medium, and > 0.5 considered large). The coefficient of determination ($R^2$) is a value that ranges from 0 to 1 and measures the proportion of the variance in the response variable that can be explained by the predictor variable in a regression model. Both coefficients were used to assess the relation between geographic distance and two different population genetic distance metrics (pairwise $F_{ST}$ and Nei's unbiased genetic distance, respectively). Z-scores were used to assess variation in landscape variables (e.g., geographic distance, quantity of softer and harder substrates) between pairwise sites within the same management unit.

## Results

### *Geographic distance has a strong, positive correlation with Nei's unbiased genetic distance*

Overall, there was a strong, positive correlation between genetic distance and geographic distance in a manner consistent with isolation-by-distance for both CCU and COU (Appendix S2.3, Supporting Information). Pearson's correlation coefficient ($r$) shows a strong, positive relation between total coastal distance and Nei's unbiased genetic distance ($r = 0.87$) and pairwise $F_{ST}$ ($r = 0.85$), respectively. The coefficient of determination ($R^2$) also shows a strong, positive relation between total coastal distance and Nei's unbiased genetic distance ($R^2 = 0.77$) and pairwise $F_{ST}$ ($R^2 = 0.71$), respectively. The difference between using Nei's unbiased genetic distance and pairwise $F_{ST}$ is small. Similar results were found for Nei's genetic distance compared to Nei's unbiased genetic distance. Nei's unbiased genetic distance (hereafter, used interchangeably as Nei's genetic distance) was used as the main population genetic distance metric for downstream analysis.

### *Heterogeneous landscapes with softer substrates may facilitate dispersal between sites*

For the MLR analysis, the two softer substrates (SP and SR) were combined into one landscape variable, and the five harder substrates (TD, FG, HC, VO, IN) were combined into a separate landscape variable. The MLR output for CCU (Table 2.2) shows heterogeneous landscapes (i.e., the interaction of ≥2 landscape variables) that include softer substrates have a statistically significant negative relation with Nei's genetic distance (MLR: Softer substrate * Harder substrate, p-value < 0.0001; Softer substrate * Number of headlands, p-value < 0.0074). Although softer substrates as an individual variable were not significant (MLR: Softer substrate, p-value = 0.9573). Harder substrates also had the highest, and an overall large, effect size ($\eta^2 = 0.2019$; $\eta^2_p = 0.3116$; $\omega^2 = 0.1973$) compared to all other independent variables, either individually or through an interaction, including softer substrates ($\eta^2 = 0.00001$; $\eta^2_p = 0.00002$; $\omega^2 = 0.00001$).

Similar results were found for the MLR output for COU (Table 2.3), except when using average kelp area as a third independent variable (MLR: Softer substrate * Harder substrate, p-value < 0.0016; Softer substrate * Average kelp area, p-value < 0.0020). Softer substrates as an individual variable were also not significant (MLR: Softer substrate, p-value = 0.8979). Harder substrates still had the highest, but an overall moderate, effect size ($\eta^2 = 0.0584$; $\eta^2_p = 0.2124$; $\omega^2 = 0.0568$) compared to all other independent variables, either individually or through an interaction, including softer substrates ($\eta^2 = 0.00002$; $\eta^2_p = 0.00011$; $\omega^2 = 0.00001$).

### *Harder substrates and kelp vegetation may limit tidewater goby dispersal*

Within the CCU, the MLR output (Table 2.2) shows harder substrates have a statistically significant positive relation with Nei's genetic distance (MLR: Harder substrate, p-value < 0.0001). The number of headlands also has a statistically significant positive relation with Nei's genetic distance, including the interaction between harder substrates and the number of headlands (MLR: Number of headlands, p-value < 0.0001; Harder substrate * Number of headlands, p-value < 0.0001).

MLR output for the COU shows harder substrates have a positive relation with Nei's genetic distance (Table 2.3), except when including average kelp area ($m^2$) as a third independent variable (MLR: Harder substrate, p-value < 0.0001). Average kelp area also has a statistically significant positive relation with Nei's genetic distance, including the interaction between harder substrates and average kelp area (MLR: Average kelp area, p-value < 0.05; Harder substrate * Average kelp area, p-value < 0.0001).

Additionally, there is a significantly higher quantity of softer and harder substrates between two adjacent sites, Santa Ynez River (hereafter, referred to as Ynez) and Jalama Creek (hereafter, referred to as Jalama), compared to any other adjacent pairwise site within the COU (Softer substrate: z-score = 2.101, p-value < 0.05; Harder substrate: z-score: 2.467, p-value < 0.05; Appendix S2.5, Supporting Information). There is also a significantly greater geographic distance between Ynez and Jalama (z-score: 2.602, p-value < 0.01) compared to any other adjacent pairwise site comparison within the COU.

### *Spatial genetic patterns between tidewater goby populations*

STRUCTURE plots ($K \geq 2$) show distinct population structures between all four management units (Figure 2.1). Within the CCU, there are also distinct population structures despite variation in sample collection years (Appendix S2.6, Supporting Information). Within the COU, there are distinct population structures, but there is also some slight complexity. For example, the STRUCTURE plots for Gaviota Creek ($K \geq 4$) do not share a similar genotypic pattern with either of its neighboring sites (Figure 2.2). Furthermore, the evidence from the STRUCTURE plots, DAPC (Figure 2.3), and the genetic distance metrics (Appendix S2.7-2.9, Supporting Information) for COU do not completely align with the current subunit designations.

For example, the STRUCTURE plots show similar genotypic patterns between two adjacent sites, Santa Ynez River and Jalama Creek, but each site is assigned to a different subunit. This is consistent with the lower Nei's unbiased genetic distance between Ynez and Jalama. Point Arguello is also a headland that separates these two sites (Figure 2.2), which is reflected in the MLR output (Appendix S2.10, Supporting Information), but this model has a much higher AIC than the MLR model using average kelp area as the third independent variable within COU (Appendix S2.4, Supporting Information).

### *The effect of temporal variation on spatial genetic variation within management units*

Data was analyzed from samples collected at two separate collection years from the same site in the CCU and SCU. STRUCTURE plots for CCU ($K > 2$) and SCU ($K = 2$) show similar genetic structures for sites despite samples collected from different years (Appendix S2.6, Supporting Information). The main exception would be one site in the CCU, Arroyo del Corral in 1990 and 2008, which shows different genotypes between the two years ($K > 2$). Arroyo del Corral was excluded from downstream analysis due to potential evidence for extirpation-recolonization, which is briefly described in the discussion. When excluding Arroyo del Corral, all three genetic distance metrics (Nei's, Nei's unbiased, and pairwise $F_{ST}$) support the evidence for genetic similarity between different years at the same site (Appendix S2.7-2.9, Supporting Information). DAPC analysis also shows no distinct clusters based on collection year within the CCU (Appendix S2.11, Supporting Information). SCU is also distinctly different from the other three management units along the LD1 axis (Appendix S2.12, Supporting Information). SCU was ignored in downstream analysis due to DAPC further demonstrating that SCU is home to a separate species, *E. kristinae* (Swift et al., 2016). It should be noted that the DAPC assigned some individuals to a management unit different from where they were originally collected (Appendix S2.13, Supporting Information). STRUCTURE HARVESTER (Earl & vonHoldt, 2012) estimated the optimum $K$ value to be $K = 2$ between and within all four management units.

### *Spatial genetic patterns between tidewater goby populations and potential implications for conservation management within the Conception management unit*

STRUCTURE plots ($K \geq 2$) show distinct population structures between all four management units (Figure 2.1). Within the CCU, there are also distinct population structures despite variation in sample collection years (Appendix S2.6, Supporting Information). Within the COU, there are distinct population structures, but there is also some slight complexity. For example, the STRUCTURE plots for Gaviota Creek ($K \geq 4$) do not share a similar genotypic pattern with either of its neighboring sites (Figure 2.2).

Furthermore, the evidence from the STRUCTURE plots, DAPC (Figure 2.3), and the genetic distance metrics (Appendix S2.7-2.9, Supporting Information) for COU do not completely align with the current subunit designations. The STRUCTURE plots ($K = 5$) and DAPC are very similar. Sites within subunit CO1 share the same genotype across all $K$ values in the STRUCTURE plots. Sites within subunit CO2 share the same genotype at $K = 2$, then reflect different population structures within the subunit at $K > 2$. Sites within subunit CO3 generally have similar genotypes at $K = 2$, then reflect more complex population structures with increasing values of $K$.

For STRUCTURE plot ($K = 2$), there seems to be a transition in genotype at Jalama, which shares a genotype with its neighboring sites, Ynez and Cañada del Cojo (hereafter, referred to as Cojo) despite Ynez being assigned to subunit CO2 and Jalama and Cojo being assigned to subunit CO3. This genetic similarity is consistent with the lower Nei's unbiased genetic distance between these three sites.

## Discussion

### *Heterogeneous landscapes with softer substrates may facilitate dispersal between sites*

Geographic distance has a slightly stronger and more positive correlation to Nei's unbiased genetic distance than with pairwise $F_{ST}$. Variation in results may be that pairwise $F_{ST}$ can become inflated by factors that reduce within-population genetic variation, such as inbreeding (Charlesworth, 1998; Nei, 1973). Thus, pairwise $F_{ST}$ may be a better indicator for inbreeding within subsamples rather than a metric for genetic differentiation between subsample pairs (Sere et al., 2017). As a result, Nei's unbiased genetic distance (hereafter, used interchangeably as Nei's genetic distance) was used for downstream analysis.

The MLR output for the Central Coast management unit (CCU) and Conception management unit (COU) supports that the interaction of different landscape variables with softer substrates (i.e., perennial sand, and sand spatially and/or temporarily mixed with rock) has a statistically significant negative correlation with Nei's genetic distance. Regardless of the other type of landscape variable used (e.g., harder substrates, number of headlands, or average area of kelp vegetation), heterogeneous landscapes (i.e., the interaction of ≥2 landscape variables) that include softer substrates suggest a genetic similarity between populations. This further implies that softer sediment may facilitate tidewater goby dispersal between sites in the presence of harder substrates, headlands, or kelp vegetation.

Softer sediment may facilitate tidewater goby dispersal since gobies are typically associated with this substrate type. For example, the reproductive behavior of tidewater gobies is moderately sex-role reversed where the male digs burrows in sand or other soft substrates to build nests (Swenson, 1997; Swift et al., 1989). Additionally, the semi-translucent olive-brown color of tidewater gobies allows them to blend well with sandy substrates (Swenson, 1997), which may enable them to hide more easily from predators during dispersal. The tidewater goby's lifestyle with sand and softer substrates may be what encourages the adults to disperse along heterogeneous landscapes that include sandy substrates.

### *Harder substrates and kelp vegetation may be resistant to tidewater goby dispersal*

The MLR output for CCU and COU shows that harder substrates have a statistically significant positive correlation with Nei's genetic distance. This suggests a homogeneous landscape of harder substrates between sites may lead to greater genetic differentiation between populations. Additionally, the MLR results show that harder substrates have the highest effect size compared to other landscape variables, either individually or through an interaction. This suggests that harder substrates may impact tidewater goby dispersal the most compared to all other landscape variables assessed.

Furthermore, the number of headlands in the CCU and the average area of kelp vegetation in the COU each respectively have a statistically significant positive correlation with Nei's genetic distance. Similar results were found for the interaction of harder substrates with each of these two variables, respectively. This suggests that headlands and kelp vegetation, each individually and their respective interaction with harder substrates, promote genetic differentiation between populations. This further implies that harder substrates, headlands, and kelp vegetation may be resistant to gene flow. This is consistent with previous studies that suggest rocky substrates and headlands potentially act as dispersal barriers to gene flow between more interactive tidewater goby populations that are separated exclusively by softer, sandy substrates (Barlow, Michele, 2002; Dawson et al., 2001a; U.S. Fish and Wildlife Services, 2005). One hypothesis is that tidewater gobies would not enter rocky areas because their light color as bottom-dwelling fish would subject them to predation in those settings. This may also be true if there is kelp vegetation between sites during dispersal; although, kelp vegetation could also offer temporary refuge from predators. An alternate hypothesis may be that there is potentially a higher abundance or diversity of predators in rocky habitats or kelp forests, which would discourage tidewater gobies from roaming these areas.

For future studies, it would be helpful to uncover details on the threshold between what constitutes a heterogeneous vs. homogeneous landscape. Based on our results, heterogeneous landscapes that include softer sediment may facilitate movement between sites while homogeneous landscapes, particularly comprised entirely of harder substrates, can be resistant to dispersal. It may be informative to identify more details on the composition of the heterogeneous landscape, such as the order in which different landscape variables appear along the coastline, which our study did not assess. This knowledge may help explain the variation observed in the population genetic structures, which is further discussed below. It should also be noted that headlands are composed of some of the rock types, specifically extrusive and intrusive volcanic rocks, that were also included in the harder substrate landscape variable for the MLR. Although avoiding this slight overlap in using harder substrates and the number of headlands as two separate variables in the MLR for the CCU analysis may be inevitable given that extrusive and intrusive volcanic rocks are not limited to forming headlands specifically.

### *Implications for how landscape variables affect tidewater goby metapopulation dynamics*

The intermittent dispersal, and consequently the metapopulation dynamics, of tidewater gobies are dependent on a sequence of required events: a heavy rainfall event followed by the simultaneous breaching of lagoons. Tidewater gobies typically live up to one year where reproduction presumably occurs during the summer months (i.e., during no to low rainfall) (Swift et al., 1989). In California, rainfall events preferentially occur during the winter months, which informed the hypothesis that tidewater gobies likely disperse as adults during the rainy season rather than having a marine larval life history stage. In fact, the mortality of smaller (i.e., younger) tidewater gobies is significantly higher with increases in salinity (Hellmair & Kinziger, 2014). Additionally, rainfall varies across the California coast, where there is typically more rain in the northern region than in the southern region. This suggests tidewater gobies in the southern region of California should presumably have lower dispersal rates and higher intermittency of dispersal than their northern counterparts. This further implies that extinction-recolonization events may be more rare in the southern regions. Tidewater goby dispersal is then also tied to the heterogeneous landscape that exists between neighboring breached lagoons. When assessing how landscape variables impact genetic structures within and between tidewater goby metapopulations, it would be important to consider other factors that can affect population structures.

Some factors that can influence population structures include geographic distance between sites and dispersal rate. Traditionally, greater distances between sites presumably promote higher genetic differentiation between populations, which is consistent with isolation-by-distance hypotheses (Wright, 1951). Simulations also suggest that historical events and landscapes could have long-term effects that confound inferences about the impacts of current landscape features on gene flow for species with very little long-distance dispersal (Landguth et al., 2010). Additionally, dispersal can affect population structures across different timescales. For example, the pattern of genetic diversity distribution among populations is dependent both on migration over geological times (historical scale dispersal) and on current migration patterns (ecological scale dispersal) (Milana et al., 2012; Miller-Sims et al., 2008). To focus more on metapopulation dynamics, we expand on the discussion related to dispersal.

In general, low dispersal or movement of organisms traditionally generates more defined population structures. While it has been hypothesized that greater dispersal rates lead to less defined population genetic structures, there has been evidence for low dispersal rates leading to low genetic differentiation (Gu et al., 2015; Morinha et al., 2017). However, dispersal in a metapopulation context can be intermittent. Tidewater gobies are considered to have low, intermittent dispersal rates since they spend most of their lifetime seasonally isolated in lagoons, restricting their dispersal to be completely dependent on rainfall events during the winter months (Dawson et al., 2001a; Swenson, 1999; Swift et al., 1989). This dispersal behavior has created extensive local genetic subdivision within and between metapopulations (Barlow, Michele, 2002; Dawson et al., 2001a; Earl et al., 2010), which has informed the designations for management units and subunits for the genus (U.S. Fish and Wildlife Services, 2005, 2014) where dispersal is thought to mostly occur within subunits only.

When assessing the STRUCTURE plots between all four management units, there are distinct population structures, which is consistent with the low dispersal rate of tidewater gobies. Within the CCU, there are also defined population structures (see Appendix S2.6, Supporting Information) that align with the current subunit designations (Dawson et al., 2001a; U.S. Fish and Wildlife Services, 2005).

Although, within the COU, there seem to be slightly more complex population structures. Some of the complexity lies in the structural patterns around sites connecting subunits CO2 and CO3, particularly across four sites starting from Santa Ynez River, Jalama Creek, Cañada del Cojo, and Gaviota Creek, in geographic order. Ynez has been described as housing a preserved and abundant tidewater goby population that likely serves as a source population during dispersal (U.S. Fish and Wildlife Services, 2005). Jalama represents a sink population that can be recolonized, which suggests the population has intermittent presence. Cojo represents a potential source population, but has a much lower abundance compared to Ynez. Gaviota is another potential source population but has variable abundance. Ynez, Jalama, and Cojo are also geographically closer to one another, whereas there are about seven other potential tidewater goby habitats separating Cojo and Gaviota (U.S. Fish and Wildlife Services, 2005).

The shared genotypes in the STRUCTURE plots and the lower Nei's genetic distance between Ynez and Jalama support the hypothesis that Ynez likely serves as a source population for Jalama, despite Cojo being slightly closer to Jalama than Ynez. There is also a significantly higher quantity of softer and harder substrates between Ynez and Jalama than any other adjacent pairwise site comparison within the COU. Per our previous discussion, this suggests the interaction of these two landscape variables likely facilitates movement between these two sites due to the presence of softer substrates. The shared genetic similarity between Ynez and Jalama also occurs in the presence of one headland, Point Arguello, that separates the two sites.

While Ynez and Jalama may be evidence that heterogeneous landscapes that include softer substrates may facilitate dispersal, despite the presence of a rocky headland, this concept becomes less clear when applied to Cañada del Cojo and Gaviota Creek. While Cojo shares a higher Nei's genetic distance with Jalama and Ynez, respectively, the STRUCTURE plots ($K = 3$

and 4) reflect similar genotypic patterns between these three sites. This provides conflicting results that Cojo may or may not be a source population for Jalama, but it seems as though there is evidence that dispersal between these two sites is possible despite the presence of one headland, Point Conception, that separates these two sites. On the other hand, Gaviota reflects a genotypic pattern that is distinct from either of its neighboring adjacent sites; however, this may be a consequence of the lack of data in this study for the seven potential tidewater goby sites that separate Cojo and Gaviota.

This specific region within COU may be evidence of the consequences of intermittent dispersal as a product of metapopulation dynamics. It becomes challenging to apply a static or uniform process to what we understand to be an intermittent, metapopulation process that is largely dictated by extirpation-recolonization dynamics rather than uniform gene flow. Assessing how landscape genetics relate to metapopulation dynamics should account for the complexities of intermittent dispersal and how this may impact the results. For example, it would be important to better characterize how population structures are influenced by the dynamic and intermittent nature of metapopulation processes, and how this relates to the frequency of sampling. Additionally, it may be challenging to tease apart actual migration events that share similar genotypic patterns that arose from shared population histories and low genetic separation. Varying rates of dispersal within metapopulations (i.e., identifying differences between and within source vs. sink sites) may also influence how we conceptualize landscape variables to affect population structures.

### *Timing of sample collections may not significantly impact population structures in the Conception management unit*

While the COU samples were collected across various years, this likely did not lead to the observed complex population structures. To assess this, we analyzed the data for sites that were collected from two different years within the CCU and South Coast management unit (SCU). When ignoring Arroyo del Corral, which may be evidence for an extirpation-recolonization event (see Chapter 1 "Hydrologic Control of Metapopulation Dispersal in the Endangered Northern Tidewater Goby (*Eucyclogobius newberryi*) on the California Coast"), all sites within CCU shared similar population genetic structures despite samples being collected from two different years. The DAPC and genetic distance metrics for all within-site pairwise comparisons in the CCU also support this. Similar results were found for SCU, although there were fewer samples analyzed from the SCU in this study. Overall, this suggests there are similar genetic structures between tidewater goby populations despite variation in collection year. This further implies that the relative connectivity between sites may not vary with respect to time. Given this evidence, we consolidated microsatellite data for the individual sites that had more than one collection year, specifically for COU, which had samples collected in 1999, 2002, and 2006.

### *Implications for conservation management within the Conception management unit*

The 2005 U.S. Fish and Wildlife Service Recovery Plan for *Eucyclogobius* partly defined the subunits within management units based on genetic evidence, but also on the premise that populations in habitats were separated by long geographic distances, especially where there was harder rocky substrate present between sites. This includes using the presence of headlands as potential indicators of separate genetic entities and, in turn, potentially separate subunits. However, our results suggest that substrate may play a more important role than headlands alone.

Within the COU, there are three main headlands: Point Sal separates Santa Maria River and Shuman Lagoon, Point Arguello separates Santa Ynez River and Jalama Creek, and Point Conception separates Jalama Creek from Cañada del Cojo. These three headlands became the main factors in designating the three subunits in COU. Santa Maria River also separates the San Luis Obispo and Santa Barbara Counties, which was additionally used as a contributing factor in defining the COU subunits (U.S. Fish and Wildlife Services, 2005). However, based on our analysis, using headlands to define the boundary locations for management subunits could be improved upon.

For example, between Santa Ynez River and Jalama Creek lies Point Arguello, which was previously used to define a subunit boundary based on mitochondrial sequencing and the presumptive influence of headlands (Dawson et al., 2001a). Ynez and Jalama were then assigned to subunits CO2 and CO3, respectively. However, based on slightly different sampling with microsatellite data, the STRUCTURE plots do not completely align with the current subunit designations (U.S. Fish and Wildlife Services, 2005). The STRUCTURE plots and Nei's genetic distance suggest the populations from Ynez and Jalama are genetically similar. There may also be evidence that dispersal is possible between Ynez and Cañada del Cojo, one additional site south of Jalama. Additionally, the STRUCTURE plots ($K > 2$) and Nei's genetic distance suggest dispersal is possible somewhere between San Luis Obispo Creek through San Antonio Creek, connecting sites from subunits CO1 and CO2. This specific example challenges both the criteria of using headlands and counties as defining boundaries between subunits. Given the subtle differences between these results and previous interpretations, some changes in how management subunits are defined may be merited. This may include using multiple sources of genetic evidence to infer subunit designations or considering that headlands in the presence of softer substrates may not completely restrict gene flow between populations as previously thought.

### *Summary*

It is important to assess how landscape variables may affect population genetic structures and dispersal, particularly how they relate to metapopulation dynamics. This can further help inform conservation management for an endangered genus like *Eucyclogobius*. Our study provides evidence that harder substrates may limit coastal dispersal of tidewater gobies and that heterogeneous landscapes that include softer substrates may facilitate movement between sites, even where some hard substrate or rocky headlands may be present. There is also evidence that harder substrates, headlands, and kelp vegetation may be resistant to gene flow. These observations should be considered in future assessments of management subunit boundaries for tidewater goby conservation.

Table 2.1: Information regarding sample collections, including latitude and longitude coordinates, collection years, number of individuals, number of microsatellite loci, average number of alleles per locus, and expected heterozygosity ($H_{exp}$) for each respective site. Sites are listed in geographic order from north to south.

| Site name | Site name abbreviation | Management unit | Subunit* | Latitude | Longitude | Collection year(s) | No. of individuals | No. of microsatellite loci | Avg no. of alleles per locus** | $H_{exp}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Arroyo del Corral | COR | Central Coast Unit (CCU) | 1 | 35°41'05 | 121°17'10 | 1990, 2008 | 42 | 14 | 4.00 | 0.420 |
| Arroyo Laguna | LAG | | 2 | 35°39'48 | 121°12'39 | 1990, 2008 | 44 | 14 | 4.57 | 0.441 |
| Arroyo de Tortuga | TOR | | 2 | 34°42'43 | 120°35'50 | 1990, 2008 | 42 | 14 | 3.79 | 0.385 |
| Arroyo del Puerto | PUE | | 2 | 35°38'35 | 121°11'15 | 1990, 2008 | 42 | 14 | 4.14 | 0.417 |
| Broken Bridge Creek | BRO | | 2 | 35°38'31 | 121°10'54 | 2008 | 30 | 14 | 3.64 | 0.430 |
| Little Pico Creek | LPC | | 2 | 35°38'02 | 121°09'46 | 1990, 2008 | 42 | 14 | 4.29 | 0.446 |
| Pico Creek | PIC | | 2 | 35°36'56 | 121°08'53 | 1990, 2008 | 42 | 14 | 4.21 | 0.449 |
| San Simeon Creek | SIM | | 2 | 35°35'43 | 121°07'35 | 1990, 2008 | 47 | 14 | 4.64 | 0.482 |
| Santa Rosa Creek | ROS | | 2 | 35°34'08 | 121°06'36 | 2008 | 30 | 14 | 3.71 | 0.535 |
| Villa Creek | VIL | | 3 | 35°32'12 | 120°53'31 | 1990, 2008 | 48 | 14 | 4.71 | 0.511 |
| San Geronimo Creek | GER | | 3 | 35°26'59 | 120°54'27 | 1990, 2008 | 43 | 14 | 4.50 | 0.463 |
| Cayucos Creek | CAY | | 3 | 35°26'58 | 120°54'26 | 1990, 2008 | 44 | 14 | 5.14 | 0.477 |
| Little Cayucos Creek | LCA | | 3 | 35°26'51 | 120°54'09 | 2008 | 30 | 14 | 4.36 | 0.461 |
| San Luis Obispo Creek | SLO | Conception Unit (COU) | 1 | 35°10'46 | 120°44'14 | 2008 | 29 | 10 | 3.40 | 0.423 |
| Mouth of Pismo Creek | PIS | | 1 | 35°08'01 | 120°38'23 | 2008 | 29 | 10 | 3.90 | 0.489 |
| Arroyo Grande | GRA | | 1 | 35°05'58 | 120°37'45 | 2008 | 29 | 10 | 3.60 | 0.477 |
| Santa Maria River | MARI | | 1 | 34°58'15 | 120°38'57 | 2008 | 30 | 10 | 3.40 | 0.467 |
| Shuman Lagoon | SHU | | 2 | 34°50'41 | 120°35'44 | 2008 | 30 | 10 | 3.30 | 0.447 |
| San Antonio Creek | ANT | | 2 | 34°48'07 | 120°37'06 | 2008 | 30 | 10 | 2.70 | 0.377 |
| Santa Ynez River | YNZ | | 2 | 34°41'31 | 120°36'03 | 2008 | 30 | 10 | 4.20 | 0.381 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Jalama Creek | JAL | | 3 | 34°30'40 | 120°30'06 | 2008 | 30 | 10 | 4.30 | 0.436 |
| Canada del Cojo | COJ | | 3 | 34°27'11 | 120°24'54 | 2006 | 30 | 10 | 2.40 | 0.291 |
| Gaviota Creek | GAV | | 3 | 34°28'15 | 120°13'25 | 1999 | 30 | 10 | 3.40 | 0.397 |
| Refugio Creek | REF | | 3 | 34°27'46 | 120°04'09 | 1999 | 30 | 10 | 2.30 | 0.285 |
| Tecolote Canyon | TEC | | 3 | 34°25'57 | 119°54'59 | 2002 | 30 | 10 | 3.20 | 0.387 |
| Arroyo Burro | BUR | | 3 | 34°25'41 | 119°44'59 | 1999 | 32 | 10 | 3.30 | 0.384 |
| Mission Creek | MIS | | 3 | 34°24'45 | 119°41'12 | 2002 | 30 | 10 | 3.20 | 0.418 |
| Andre Clark Bird Refuge | AND | | 3 | 34°25'00 | 119°39'47 | 2002 | 30 | 10 | 3.30 | 0.410 |
| Arroyo Paredon | PAR | | 3 | 34°24'49 | 119°33'33 | 2002 | 30 | 10 | 3.50 | 0.443 |
| Carpinteria Creek | CAR | | 3 | 34°23'25 | 119°31'10 | 1999 | 30 | 10 | 2.90 | 0.368 |
| Rincon Creek | RIN | | 3 | 34˚22'25 | 119˚28'37 | 2002 | 30 | 10 | 2.20 | 0.237 |
| Ventura River | VEN | Los Angeles/ Ventura Unit (LAVU) | 1 | 34°16'32 | 119°18'25 | 1999 | 30 | 11 | 2.57 | 0.327 |
| Santa Clara River | CLA | | 1 | 34°14'07 | 119°15'46 | 1999 | 25 | 11 | 3.07 | 0.413 |
| Ormond Beach | ORM | | 1 | 34°08'13 | 119°11'00 | 1999 | 30 | 11 | 2.93 | 0.426 |
| Malibu Beach | MAL | | 1 | 34°01'58 | 118°40'48 | 1999 | 29 | 11 | 2.64 | 0.387 |
| Topanga Creek | TOP | | 1 | 34°02'19 | 118°34'50 | 2001 | 27 | 11 | 2.07 | 0.327 |
| San Onofre Creek | ONO | South Coast Unit (SCU) | 1 | 33°22'49 | 117°34'39 | 1990, 2007 | 40 | 6 | 1.50 | 0.123 |
| Las Flores Creek | FLO | | 2 | 33°17'26 | 117°27'50 | 1990, 2007 | 51 | 6 | 1.86 | 0.155 |
| Aliso Canyon | ALI | | 2 | 33°15'54 | 117°26'31 | 2007 | 30 | 6 | 1.36 | 0.147 |
| Santa Margarita River | MARG | | 2 | 33°13'55 | 117°24'55 | 1990, 2007 | 19 | 6 | 1.57 | 0.148 |

*Current designation defined by the 2005 U.S. Fish and Wildlife Service Recovery Plan
**Based on combined collection years if more than one

Table 2.2: Multiple linear regression (MLR) model output for dependent variable (Nei's unbiased genetic distance) and independent variables (Softer substrates, harder substrates, and number of headlands) for the CCU dataset. This MLR output had the lowest AIC compared to another model using the average kelp area instead of the number of headlands as a third independent variable (see Appendix S2.4, Supporting Information). Asterisk represents the interaction between two or more variables. Outputs with significant p-values are in bold.

| **Variable** | **Estimate** | **Std. Error** | **P-value** | **Eta squared** ($\eta^2$) | **Partial eta squared** ($\eta^2_p$) | **Omega squared** ($\omega^2$) |
|---|---|---|---|---|---|---|
| Softer substrates | -8.362e-8 | 1.557e-6 | 0.9573 | 0.00001 | 0.00002 | 0.00001 |
| **Harder substrates** | **5.0606e-6** | **6.984e-7** | **<0.0001** | 0.20186 | 0.31156 | 0.19726 |
| **Number of headlands** | **0.04527** | **0.011039** | **<0.0001** | 0.06467 | 0.12662 | 0.06059 |
| **Softer substrates * Harder substrates** | **-6.88e-10** | **1.48e-10** | **<0.0001** | 0.08340 | 0.15753 | 0.07925 |
| **Softer substrates * Number of headlands** | **-4.719e-6** | **1.732e-6** | **0.0074** | 0.02854 | 0.06013 | 0.02460 |
| **Harder substrates * Number of headlands** | **6.9954e-6** | **1.411e-6** | **<0.0001** | 0.09057 | 0.16878 | 0.08639 |

Table 2.3. Multiple linear regression (MLR) model output for dependent variable (Nei's unbiased genetic distance) and independent variables (Softer substrates, harder substrates, and average area of kelp) for the COU dataset. This MLR output had the lowest AIC compared to another model using the number of headlands instead of the average kelp area as a third independent variable (see Appendix S2.4, Supporting Information). Asterisk represents the interaction between two or more variables. Outputs with significant p-values are in bold.

| **Variable** | **Estimate** | **Std. Error** | **P-value** | **Eta squared** ($\eta^2$) | **Partial eta squared** ($\eta^2_p$) | **Omega squared** ($\omega^2$) |
|---|---|---|---|---|---|---|
| Softer substrates | -1.468e-7 | 1.141e-6 | 0.8979 | 0.00002 | 0.00011 | 0.00001 |
| **Harder substrates** | **1.464e-5** | **2.333e-6** | **<0.0001** | 0.05840 | 0.21244 | 0.05683 |
| **Softer substrates * Harder substrates** | **-1.77e-10** | **5.51e-11** | **0.0016** | 0.01532 | 0.06608 | 0.01382 |
| **Avg kelp area ($m^2$)** | **3.6917e-7** | **1.803e-7** | **0.0424** | 0.00621 | 0.02790 | 0.00472 |
| **Softer substrates * Avg kelp area ($m^2$)** | **-8.33e-12** | **2.64e-12** | **0.0020** | 0.01475 | 0.06379 | 0.01325 |
| **Harder substrates * Avg kelp area ($m^2$)** | **6.742e-11** | **1.53e-11** | **<0.0001** | 0.02864 | 0.11882 | 0.02711 |

Figure 2.1: STRUCTURE plots (*K* ≥ 2) for all four management units: Central Coast Unit (CCU), Conception Unit (COU), Los Angeles/Ventura Unit (LAVU), and South Coast Unit (SCU). Management units with samples collected in more than one collection year have been combined (CCU and SCU) (see Table 2.1). All management units show distinct population genetic structures from one another (*K* = 4). There is also evidence of population subdivision within the COU near Santa Ynez River (YNZ) and Jalama Creek (JAL).

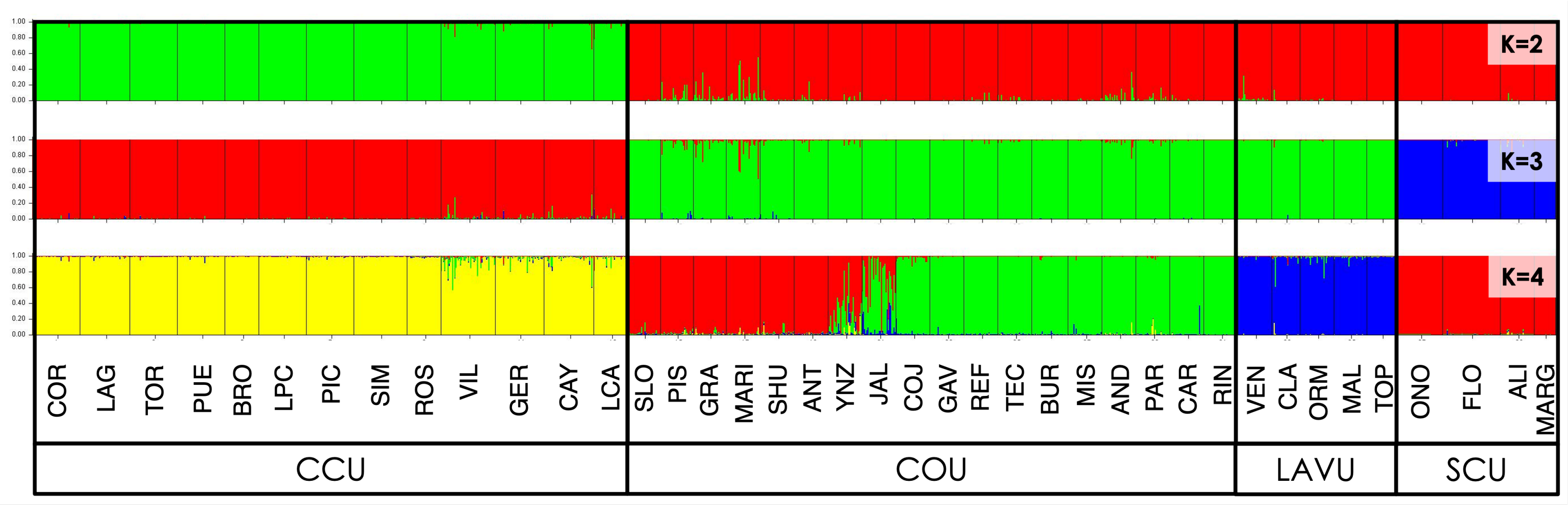


Figure 2.2: STRUCTURE plots and coast map of collection sites from the Conception management unit (COU). Asterisks denote different headlands: Point Sal separates Santa Maria River (MARI) and Shuman Lagoon (SHU); Point Arguello separates Santa Ynez River (YNZ) and Jalama Creek (JAL); and Point Conception separates JAL from Cañada del Cojo (COJ). MARI also serves as a boundary between San Luis Obispo and Santa Barbara Counties. The 2005 U.S. Fish and Wildlife Service Recovery Plan used headlands and county boundaries to define subunits within COU (CO1, CO2, and CO3); however, the genotype patterns in the STRUCTURE plots (*K* > 2) contradict the current subunit designations. Specifically, YNZ, JAL, and COJ share similar genotypic patterns, yet are assigned to different subunits. This further contradicts the assumption that dispersal mostly occurs within subunits.

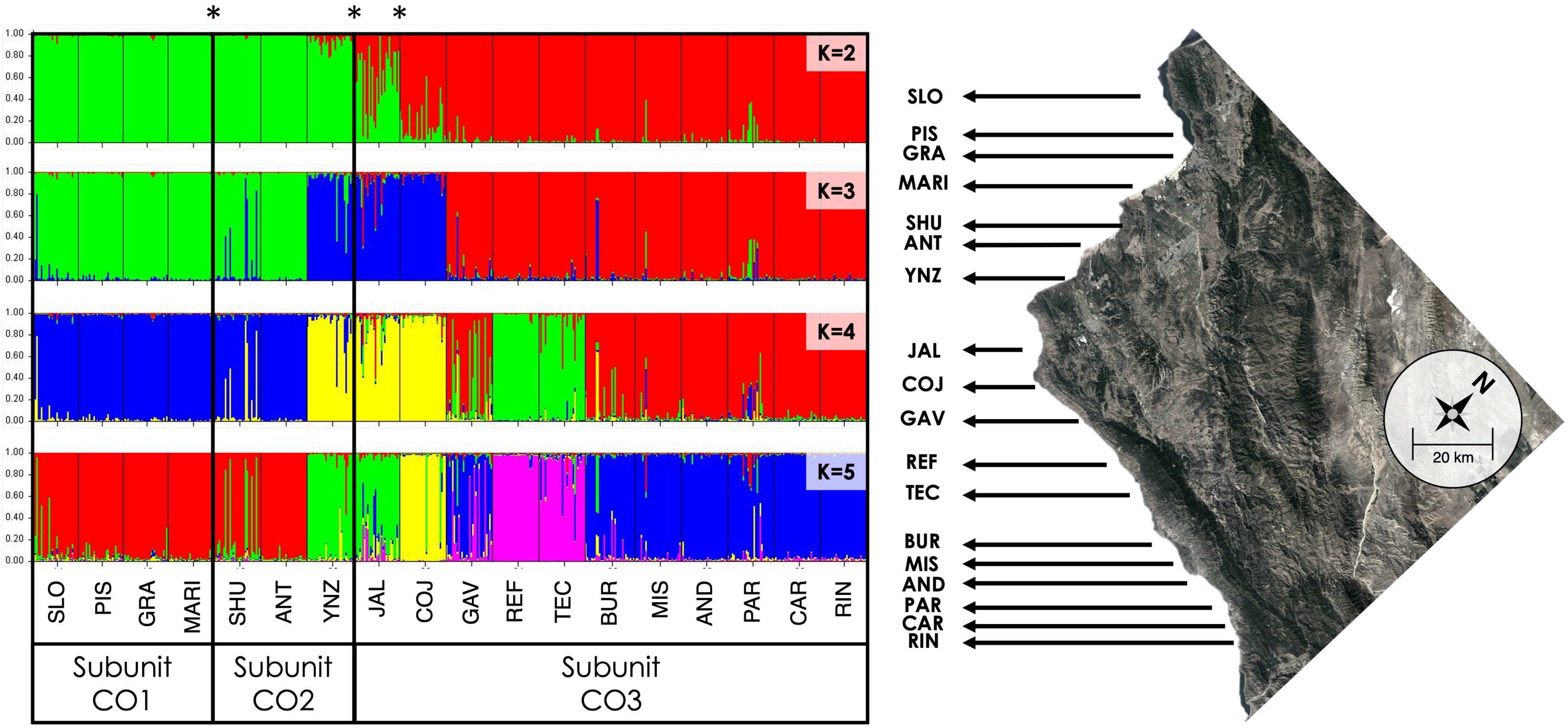
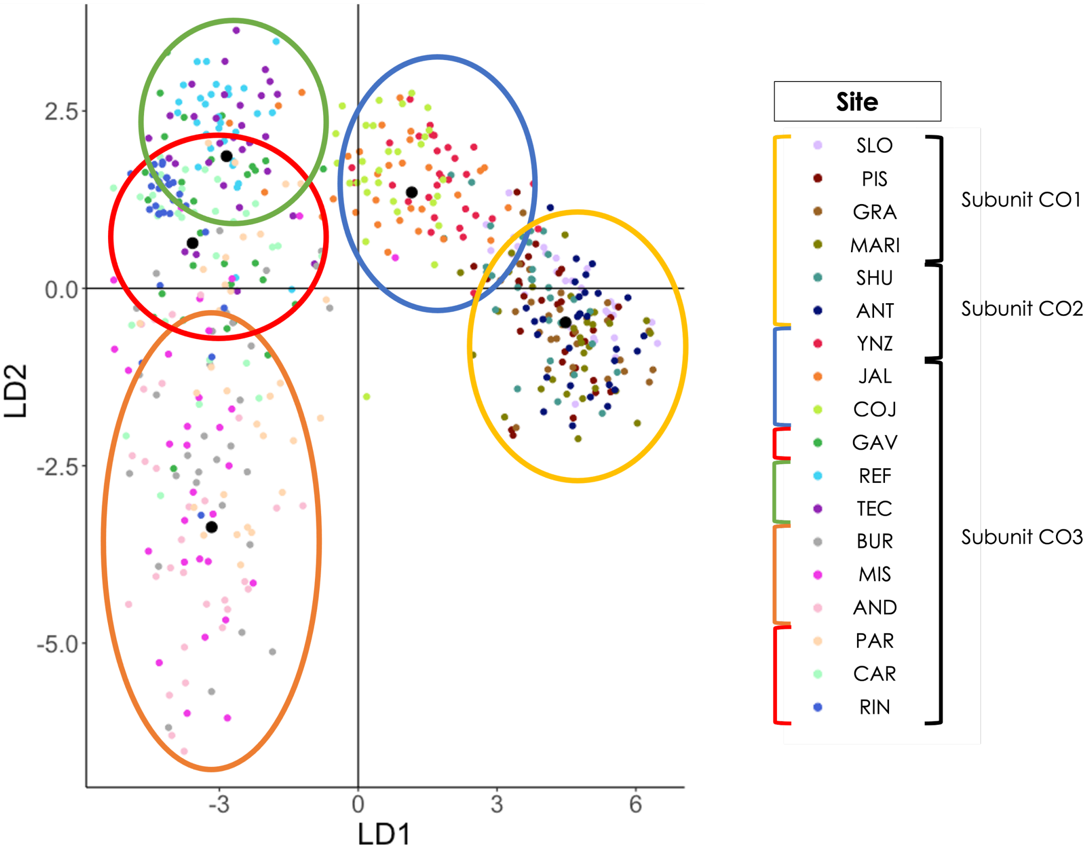


Figure 2.3: Discriminant analysis of principal components (DAPC) for the Conception management unit. Overall, the DAPC clusters individuals in a similar way to the STRUCTURE plot ($K = 5$), except Arroyo Burro (BUR), Mission Creek (MIS), and Andre Clark Bird Refuge (AND) are grouped separately from Arroyo Paredon (PAR), Carpinteria Creek (CAR) and Rincon Creek (RIN) (see Figure 2.2). Similar to the STRUCTURE plots, the assigned DAPC clusters ($K = 5$) contradict the current subunit designations defined by the 2005 U.S. Fish and Wildlife Service Recovery Plan (black brackets).



Appendix S2.1: Microsatellite loci for *Eucyclogobius*, including primer sequence (forward primer listed above reverse), length, cloned sequence repeat motif, size of range of alleles (base pairs), number of alleles observed.

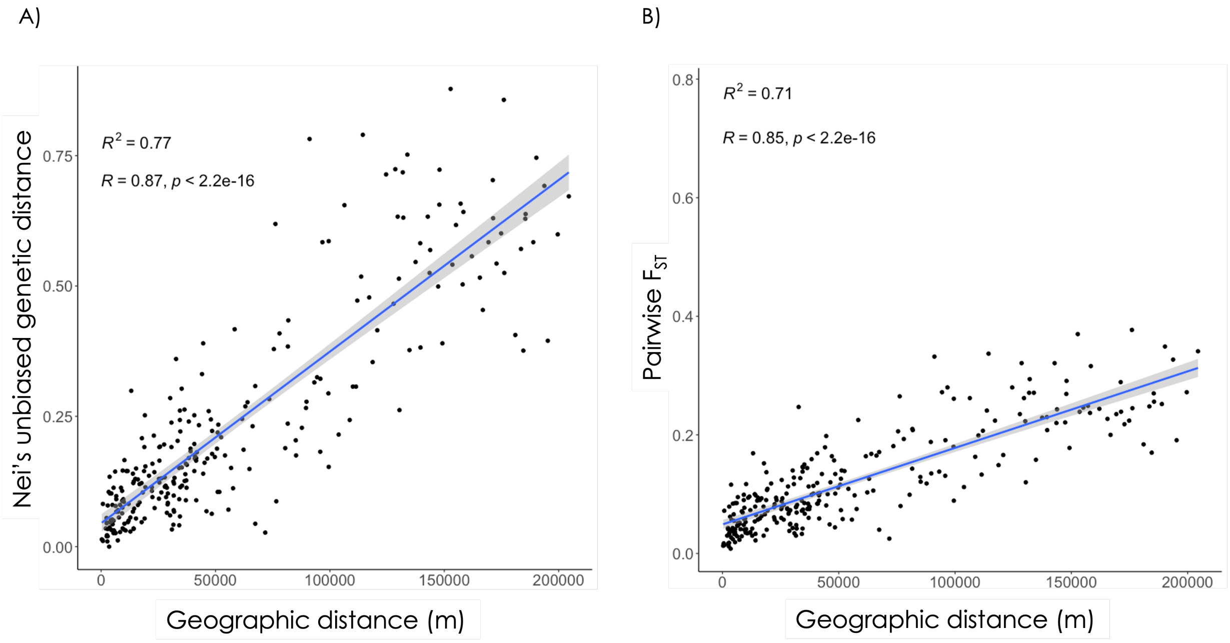
| **Locus** | **Primer sequence (5' - 3')** | **Length** | **Motif** | **Fragment size (bp)** | **No. alleles** |
|---|---|---|---|---|---|
| ENE2 | GTCGACTGGCAGTATGGGAT | 20 | $(ATCTCT)_5$ | 142-154 | 3 |
| | AGACTCAAATATGTGCACACCAC | 13 | | | |
| ENE3 | CTAACCTGCCTTGACCCAAA | 20 | $(GT)_{12}$ | 120-130 | 2 |
| | GCAGAATCAACATAAGAACTATGGAAA | 27 | | | |
| ENE5 | GCTTGTGCAGTATGGGATCTC[a] | 21 | $(GT)_4(AT)(GT)_9$ | 306–326 | 5 |
| | CTCGGAGCGTTCATTTATCTC | 21 | | | |
| ENE6 | TCAGGTTTGTGCTAAAATGATG[a] | 22 | $(CA)_{11}(CATACA)_5(CA)_4$ | 241–257 | 10 |
| | TCCGATGACCACTTGTCC | 18 | | | |
| ENE7 | TCACATGAATCGGAGACAGT[a] | 20 | $(CAT)_7(CAC)(CAT)$ | 135–159 | 6 |
| | CAGAGAGGGCACTTTTTCAG | 20 | | | |
| ENE8 | GAGGAAGGCGAGCTGATTA[a] | 19 | (ACCATCATC)4(ACC)(ATC)8 | 101–204 | 14 |
| | CGGAGAGAAGGTGTTGAGAG | 20 | | | |
| ENE9 | CCTTCATTTTTCCATCAGAAGCG[a] | 23 | $(ATG)_{28}$ | 131–209 | 27 |
| | CCTTATTTACATCTTCCCTCCA | 22 | | | |
| ENE12 | CTGGGATTGTCTTGGAACAG[a] | 20 | $(GAT)_9$ | 183–240 | 13 |
| | GGGTGTGTGTGAGAGAGTGG | 20 | | | |
| ENE15 | CCCGGAGGAGTTAGAGGAA[a] | 19 | $(TGGA)_7$ | 281–293 | 3 |
| | GAGCCTGTGGTTTGTCGAG | 19 | | | |
| ENE16 | GTCGCCTTGATTTTATTGTGA[a] | 21 | $(TGGA)_6$ | 138–216 | 6 |
| | CTCAGCGTGGTTTCATTAT | 19 | | | |
| ENE17 | CAGAGGTAGATCAGAAGAAC[a] | 20 | $(ATCC)_6$ | 165–173 | 2 |
| | CCGGATAAAGTGCAGAAAAT | 20 | | | |
| ENE18 | GGAGAACGAGAGAGAAAGA[a] | 19 | $(GA)_4(AC)(GA)_{25}$ | 132–152 | 8 |
| | GGCTGGTGTTTGATACATC | 19 | | | |
| ENE19 | CGCGTCAGTTTTCACCTTTA[a] | 20 | $(TCTA)_{11}$ | 110–134 | 3 |

| | | | | | |
|---|---|---|---|---|---|
| | GAGAATGCCCAAAATCACC | 19 | | | |
| ENE21 | TGCAGAGAAAGAGACAGGTATT[a] | 22 | $(CATCAA)_2(CAT)_7(CTG)_2(CAT)_3$ | 154–168 | 3 |
| | ATTGAGGTGCTGACACTGAG | 20 | | | |

[a] 23 nucleotide M13 sequence (5'-AGGGTTTTCCCAGTCACGACGTT-3') added to the 50 end of the forward primer to allow annealing of the dye-labeled M13 primer to the PCR product

Appendix S2.2: Examples of headlands (white arrows) identified between pairwise sites based on visual estimation from satellite imagery via Google Earth Pro (Landsat 7). One headland is between Santa Maria River and Shuman Lagoon (A). Two headlands are between Tecolote Canyon and Arroyo Burro (B). There are no headlands between Arroyo Burro and Mission Creek (C) or between Gaviota Creek and Refugio Creek (D).

Appendix S2.3: Plot showing the relation between geographic distance vs Nei's unbiased genetic distance (A) and Pairwise $F_{ST}$ (B). Geographic distance has a slightly stronger relation with Nei's unbiased genetic distance. Data includes CCU and COU.



Appendix S2.4: AIC values for different MLR model outputs using Nei's unbiased genetic distance as the dependent variable and various landscape variables as independent variables (e.g., softer substrates, harder substrates, number of headlands, average kelp area). The respective models with lower AIC values for CCU and COU are in bold.

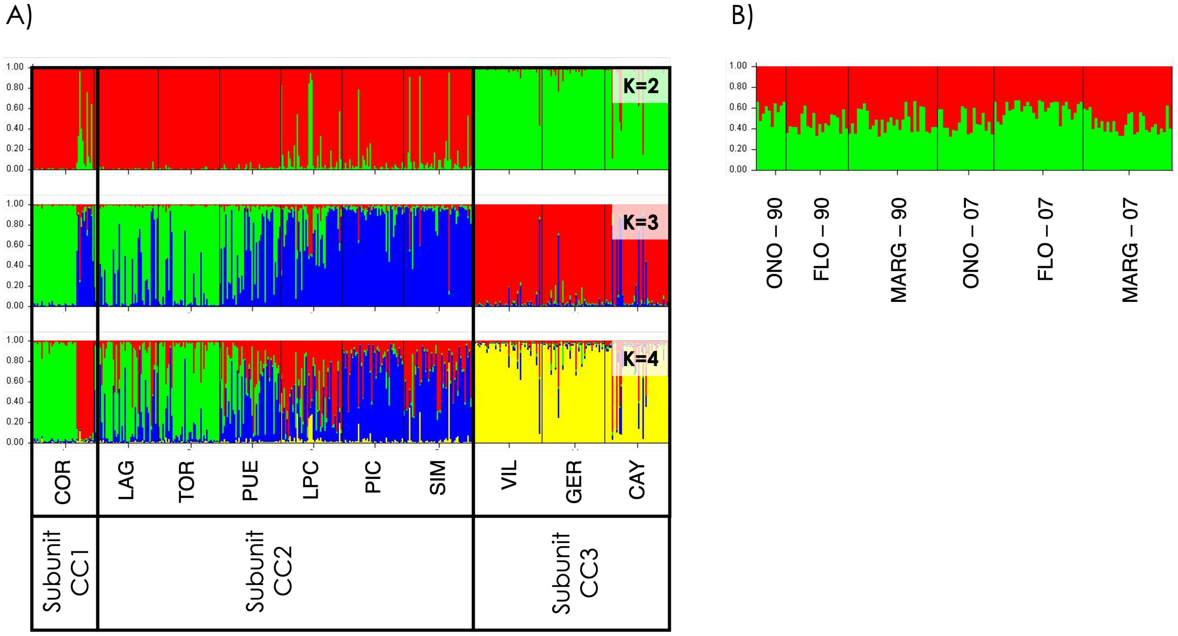
| **Dataset** | **Softer substrates** | **Harder substrates** | **Number of headlands** | **Avg kelp area ($m^2$)** | **AIC** |
|---|---|---|---|---|---|
| CCU | x | x | | x | -365.79 |
| **CCU** | **x** | **x** | **x** | | **-384.411** |
| **COU** | **x** | **x** | | **x** | **-233.929** |
| COU | x | x | x | | -214.279 |

Appendix S2.5: Z-scores for landscape variables between adjacent pairwise sites in COU. Significant z-scores are highlighted in yellow.

| Main Site | Collection Year | Management Unit | Paired Site Name | Total Coastal Distance (m) | Total Coastal Distance (z-score) | Softer substrates (m) | Softer substrates (z-score) | Harder substrates (m) | Harder substrates (z-score) |
|---|---|---|---|---|---|---|---|---|---|
| San Luis Obispo | 2008 | COU | Pismo | 10699 | -0.1825 | 8559.2 | -0.2855 | 2139.8 | 0.0803 |
| Mouth of Pismo Creek | 2008 | COU | Grande | 3499 | -1.1732 | 3499 | -1.2075 | 0 | -0.6349 |
| Arroyo Grande | 2008 | COU | Santa Maria | 14206 | 0.3000 | 14206 | 0.7433 | 0 | -0.6349 |
| Santa Maria River | 2008 | COU | Shuman | 17658 | 0.7750 | 8829 | -0.2363 | 8829 | 2.3167 |
| Shuman Lagoon | 2008 | COU | San Antonio | 5616 | -0.8819 | 5616 | -0.8217 | 0 | -0.6349 |
| San Antonio Creek | 2008 | COU | Santa Ynez | 13156 | 0.1555 | 8551.4 | -0.2869 | 4604.6 | 0.9044 |
| Santa Ynez River | 2008 | COU | Jalama | 30938 | 2.6022 | 21656.6 | 2.1009 | 9281.4 | 2.4679 |
| Jalama Creek | 2008 | COU | Canada del Cojo | 14271 | 0.3089 | 11416.8 | 0.2351 | 2854.2 | 0.3192 |
| Canada del Cojo | 2006 | COU | Gaviota | 17825 | 0.7979 | 16042.5 | 1.0780 | 1782.5 | -0.0390 |
| Gaviota | 1999 | COU | Refugio | 14927 | 0.3992 | 14927 | 0.8747 | 0 | -0.6349 |
| Refugio Creek | 1999 | COU | Tecolote | 15232 | 0.4412 | 13708.8 | 0.6528 | 1523.2 | -0.1257 |
| Tecolote Canyon | 2002 | COU | Burro | 18208 | 0.8506 | 17297.6 | 1.3067 | 910.4 | -0.3306 |
| Arroyo Burro | 1999 | COU | Mission | 7284 | -0.6524 | 6919.8 | -0.5842 | 364.2 | -0.5132 |
| Mission Creek | 2002 | COU | Andre Clark Bird | 2023 | -1.3763 | 2023 | -1.4764 | 0 | -0.6349 |
| Andre Clark Bird Refuge | 2002 | COU | Paredon | 9705 | -0.3192 | 9705 | -0.0767 | 0 | -0.6349 |
| Arroyo Paredon | 2002 | COU | Carpinteria | 4354 | -1.0555 | 4354 | -1.0517 | 0 | -0.6349 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Carpinteria Creek | 1999 | COU | Rincon | 4832 | -0.9897 | 4832 | -0.9646 | 0 | -0.6349 |
| | | | Mean | 12025.47 | | 10126.10 | | 1899.37 | |
| | | | Stdev | 7267.64 | | 5488.17 | | 2991.16 | |

Appendix S2.6: STRUCTURE plots for 1990 and 2008 CCU samples (A) and 1990 and 2007 SCU samples (B). (A) shows the combined 1990 and 2008 CCU samples whereas (B) shows the 1990 and 2007 SCU samples separated. Unlike CCU, SCU had fewer sites for comparison between collection years (see Table 2.1). With the exception of Arroyo del Corral (COR) in the CCU, all other sites sampled across two collection years show similar genotypic population structures suggesting time may not significantly impact population genetic structures.



Appendix S2.7: Pairwise Nei's unbiased genetic distance for CCU (top) and COU (bottom).

| | COR_08 | COR_90 | LAG_08 | LAG_90 | TOR_08 | TOR_90 | PUE_08 | PUE_90 | BRO_08 | LPC_08 | LPC_90 | PIC_08 | PIC_90 | SIM_08 | SIM_90 | ROS_08 | VIL_08 | VIL_90 | GER_08 | GER_90 | CAY_08 | CAY_90 | LCA_08 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| COR_08 | 0.000 | | | | | | | | | | | | | | | | | | | | | | | COR_08 |
| COR_90 | 0.196 | 0.000 | | | | | | | | | | | | | | | | | | | | | | COR_90 |
| LAG_08 | 0.035 | 0.160 | 0.000 | | | | | | | | | | | | | | | | | | | | | LAG_08 |
| LAG_90 | 0.061 | 0.146 | 0.022 | 0.000 | | | | | | | | | | | | | | | | | | | | LAG_90 |
| TOR_08 | 0.044 | 0.153 | 0.011 | 0.023 | 0.000 | | | | | | | | | | | | | | | | | | | TOR_08 |
| TOR_90 | 0.052 | 0.166 | 0.041 | 0.082 | 0.015 | 0.000 | | | | | | | | | | | | | | | | | | TOR_90 |
| PUE_08 | 0.067 | 0.122 | 0.032 | 0.034 | 0.037 | 0.042 | 0.000 | | | | | | | | | | | | | | | | | PUE_08 |
| PUE_90 | 0.168 | 0.140 | 0.106 | 0.102 | 0.127 | 0.134 | 0.052 | 0.000 | | | | | | | | | | | | | | | | PUE_90 |
| BRO_08 | 0.073 | 0.129 | 0.043 | 0.063 | 0.051 | 0.055 | 0.013 | 0.076 | 0.000 | | | | | | | | | | | | | | | BRO_08 |
| LPC_08 | 0.120 | 0.094 | 0.070 | 0.050 | 0.067 | 0.092 | 0.034 | 0.065 | 0.055 | 0.000 | | | | | | | | | | | | | | LPC_08 |
| LPC_90 | 0.138 | 0.085 | 0.089 | 0.067 | 0.078 | 0.097 | 0.042 | 0.079 | 0.063 | 0.001 | 0.000 | | | | | | | | | | | | | LPC_90 |
| PIC_08 | 0.154 | 0.108 | 0.086 | 0.039 | 0.064 | 0.112 | 0.050 | 0.085 | 0.080 | 0.022 | 0.025 | 0.000 | | | | | | | | | | | | PIC_08 |
| PIC_90 | 0.145 | 0.114 | 0.088 | 0.033 | 0.061 | 0.107 | 0.058 | 0.113 | 0.089 | 0.055 | 0.046 | 0.004 | 0.000 | | | | | | | | | | | PIC_90 |
| SIM_08 | 0.107 | 0.115 | 0.056 | 0.024 | 0.050 | 0.099 | 0.028 | 0.090 | 0.056 | 0.026 | 0.030 | 0.020 | 0.035 | 0.000 | | | | | | | | | | SIM_08 |
| SIM_90 | 0.108 | 0.114 | 0.060 | 0.029 | 0.049 | 0.082 | 0.030 | 0.080 | 0.053 | 0.023 | 0.031 | 0.023 | 0.034 | 0.003 | 0.000 | | | | | | | | | SIM_90 |
| ROS_08 | 0.165 | 0.200 | 0.129 | 0.127 | 0.150 | 0.212 | 0.145 | 0.164 | 0.145 | 0.113 | 0.145 | 0.144 | 0.184 | 0.105 | 0.097 | 0.000 | | | | | | | | ROS_08 |
| VIL_08 | 0.390 | 0.187 | 0.303 | 0.231 | 0.261 | 0.297 | 0.238 | 0.258 | 0.285 | 0.133 | 0.155 | 0.130 | 0.142 | 0.159 | 0.140 | 0.252 | 0.000 | | | | | | | VIL_08 |
| VIL_90 | 0.291 | 0.140 | 0.220 | 0.152 | 0.199 | 0.243 | 0.160 | 0.191 | 0.207 | 0.088 | 0.115 | 0.084 | 0.092 | 0.113 | 0.095 | 0.246 | 0.033 | 0.000 | | | | | | VIL_90 |
| GER_08 | 0.222 | 0.163 | 0.183 | 0.152 | 0.170 | 0.189 | 0.114 | 0.156 | 0.151 | 0.070 | 0.083 | 0.084 | 0.108 | 0.104 | 0.104 | 0.213 | 0.102 | 0.042 | 0.000 | | | | | GER_08 |
| GER_90 | 0.293 | 0.178 | 0.218 | 0.188 | 0.211 | 0.240 | 0.150 | 0.221 | 0.166 | 0.097 | 0.098 | 0.096 | 0.102 | 0.129 | 0.131 | 0.255 | 0.112 | 0.047 | 0.029 | 0.000 | | | | GER_90 |
| CAY_08 | 0.219 | 0.094 | 0.165 | 0.117 | 0.145 | 0.165 | 0.089 | 0.099 | 0.126 | 0.041 | 0.049 | 0.042 | 0.068 | 0.065 | 0.067 | 0.191 | 0.078 | 0.024 | 0.020 | 0.039 | 0.000 | | | CAY_08 |
| CAY_90 | 0.286 | 0.105 | 0.231 | 0.170 | 0.207 | 0.252 | 0.140 | 0.159 | 0.180 | 0.072 | 0.091 | 0.090 | 0.125 | 0.106 | 0.113 | 0.211 | 0.079 | 0.026 | 0.022 | 0.045 | 0.008 | 0.000 | | CAY_90 |
| LCA_08 | 0.234 | 0.137 | 0.181 | 0.142 | 0.173 | 0.222 | 0.113 | 0.138 | 0.139 | 0.058 | 0.078 | 0.067 | 0.090 | 0.081 | 0.084 | 0.192 | 0.091 | 0.032 | 0.010 | 0.022 | 0.014 | 0.007 | 0.000 | LCA_08 |

| | SLO_08 | PIS_08 | GRA_08 | MARI_08 | SHU_08 | ANT_08 | YNZ_08 | JAL_08 | COJ_06 | GAV_99 | REF_99 | TEC_02 | BUR_99 | MIS_02 | AND_02 | PAR_02 | CAR_99 | RIN_02 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SLO_08 | 0.000 | | | | | | | | | | | | | | | | | | SLO_08 |
| PIS_08 | 0.031 | 0.000 | | | | | | | | | | | | | | | | | PIS_08 |
| GRA_08 | 0.035 | 0.000 | 0.000 | | | | | | | | | | | | | | | | GRA_08 |
| MARI_08 | 0.095 | 0.047 | 0.034 | 0.000 | | | | | | | | | | | | | | | MARI_08 |
| SHU_08 | 0.068 | 0.117 | 0.121 | 0.122 | 0.000 | | | | | | | | | | | | | | SHU_08 |
| ANT_08 | 0.105 | 0.169 | 0.157 | 0.143 | 0.031 | 0.000 | | | | | | | | | | | | | ANT_08 |
| YNZ_08 | 0.149 | 0.175 | 0.188 | 0.263 | 0.145 | 0.299 | 0.000 | | | | | | | | | | | | YNZ_08 |
| JAL_08 | 0.182 | 0.204 | 0.236 | 0.308 | 0.172 | 0.331 | 0.033 | 0.000 | | | | | | | | | | | JAL_08 |
| COJ_06 | 0.307 | 0.294 | 0.322 | 0.384 | 0.277 | 0.417 | 0.161 | 0.125 | 0.000 | | | | | | | | | | COJ_06 |
| GAV_99 | 0.466 | 0.478 | 0.518 | 0.586 | 0.434 | 0.619 | 0.269 | 0.132 | 0.208 | 0.000 | | | | | | | | | GAV_99 |
| REF_99 | 0.633 | 0.631 | 0.724 | 0.790 | 0.584 | 0.782 | 0.409 | 0.260 | 0.360 | 0.130 | 0.000 | | | | | | | | REF_99 |
| TEC_02 | 0.503 | 0.499 | 0.569 | 0.633 | 0.472 | 0.655 | 0.315 | 0.186 | 0.244 | 0.060 | 0.054 | 0.000 | | | | | | | TEC_02 |
| BUR_99 | 0.525 | 0.516 | 0.557 | 0.656 | 0.514 | 0.714 | 0.307 | 0.189 | 0.231 | 0.063 | 0.203 | 0.104 | 0.000 | | | | | | BUR_99 |
| MIS_02 | 0.571 | 0.543 | 0.584 | 0.617 | 0.546 | 0.718 | 0.354 | 0.228 | 0.283 | 0.071 | 0.197 | 0.123 | 0.012 | 0.000 | | | | | MIS_02 |
| AND_02 | 0.638 | 0.601 | 0.630 | 0.658 | 0.582 | 0.752 | 0.415 | 0.278 | 0.379 | 0.112 | 0.233 | 0.196 | 0.075 | 0.035 | 0.000 | | | | AND_02 |
| PAR_02 | 0.395 | 0.376 | 0.406 | 0.454 | 0.390 | 0.525 | 0.262 | 0.153 | 0.175 | 0.044 | 0.210 | 0.127 | 0.054 | 0.047 | 0.091 | 0.000 | | | PAR_02 |
| CAR_99 | 0.599 | 0.584 | 0.629 | 0.703 | 0.541 | 0.723 | 0.377 | 0.215 | 0.266 | 0.027 | 0.151 | 0.089 | 0.074 | 0.080 | 0.122 | 0.046 | 0.000 | | CAR_99 |
| RIN_02 | 0.672 | 0.692 | 0.746 | 0.857 | 0.642 | 0.878 | 0.382 | 0.243 | 0.325 | 0.087 | 0.245 | 0.169 | 0.117 | 0.136 | 0.191 | 0.088 | 0.052 | 0.000 | RIN_02 |

Appendix S2.8: Pairwise Nei's genetic distance for CCU (top) and COU (bottom).

| | COR_08 | COR_90 | LAG_08 | LAG_90 | TOR_08 | TOR_90 | PUE_08 | PUE_90 | BRO_08 | LPC_08 | LPC_90 | PIC_08 | PIC_90 | SIM_08 | SIM_90 | ROS_08 | VIL_08 | VIL_90 | GER_08 | GER_90 | CAY_08 | CAY_90 | LCA_08 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| COR_08 | 0.000 | | | | | | | | | | | | | | | | | | | | | | | COR_08 |
| COR_90 | 0.220 | 0.000 | | | | | | | | | | | | | | | | | | | | | | COR_90 |
| LAG_08 | 0.046 | 0.186 | 0.000 | | | | | | | | | | | | | | | | | | | | | LAG_08 |
| LAG_90 | 0.081 | 0.181 | 0.044 | 0.000 | | | | | | | | | | | | | | | | | | | | LAG_90 |
| TOR_08 | 0.054 | 0.178 | 0.023 | 0.044 | 0.000 | | | | | | | | | | | | | | | | | | | TOR_08 |
| TOR_90 | 0.067 | 0.196 | 0.057 | 0.107 | 0.031 | 0.000 | | | | | | | | | | | | | | | | | | TOR_90 |
| PUE_08 | 0.078 | 0.148 | 0.044 | 0.056 | 0.049 | 0.058 | 0.000 | | | | | | | | | | | | | | | | | PUE_08 |
| PUE_90 | 0.184 | 0.172 | 0.124 | 0.129 | 0.144 | 0.156 | 0.070 | 0.000 | | | | | | | | | | | | | | | | PUE_90 |
| BRO_08 | 0.083 | 0.155 | 0.056 | 0.085 | 0.063 | 0.071 | 0.026 | 0.094 | 0.000 | | | | | | | | | | | | | | | BRO_08 |
| LPC_08 | 0.131 | 0.121 | 0.083 | 0.072 | 0.080 | 0.109 | 0.047 | 0.084 | 0.069 | 0.000 | | | | | | | | | | | | | | LPC_08 |
| LPC_90 | 0.158 | 0.121 | 0.111 | 0.098 | 0.100 | 0.122 | 0.063 | 0.107 | 0.085 | 0.024 | 0.000 | | | | | | | | | | | | | LPC_90 |
| PIC_08 | 0.165 | 0.135 | 0.099 | 0.062 | 0.077 | 0.129 | 0.063 | 0.104 | 0.093 | 0.036 | 0.048 | 0.000 | | | | | | | | | | | | PIC_08 |
| PIC_90 | 0.165 | 0.149 | 0.110 | 0.064 | 0.082 | 0.133 | 0.080 | 0.141 | 0.111 | 0.077 | 0.078 | 0.027 | 0.000 | | | | | | | | | | | PIC_90 |
| SIM_08 | 0.119 | 0.143 | 0.070 | 0.047 | 0.063 | 0.117 | 0.042 | 0.110 | 0.070 | 0.041 | 0.054 | 0.035 | 0.058 | 0.000 | | | | | | | | | | SIM_08 |
| SIM_90 | 0.125 | 0.148 | 0.080 | 0.058 | 0.068 | 0.105 | 0.050 | 0.105 | 0.073 | 0.044 | 0.060 | 0.043 | 0.063 | 0.025 | 0.000 | | | | | | | | | SIM_90 |
| ROS_08 | 0.179 | 0.229 | 0.145 | 0.152 | 0.165 | 0.232 | 0.161 | 0.186 | 0.161 | 0.130 | 0.170 | 0.161 | 0.209 | 0.123 | 0.120 | 0.000 | | | | | | | | ROS_08 |
| VIL_08 | 0.403 | 0.215 | 0.318 | 0.255 | 0.275 | 0.316 | 0.253 | 0.279 | 0.300 | 0.148 | 0.179 | 0.145 | 0.166 | 0.175 | 0.162 | 0.271 | 0.000 | | | | | | | VIL_08 |
| VIL_90 | 0.310 | 0.173 | 0.240 | 0.181 | 0.218 | 0.267 | 0.180 | 0.217 | 0.227 | 0.108 | 0.145 | 0.105 | 0.121 | 0.135 | 0.123 | 0.270 | 0.056 | 0.000 | | | | | | VIL_90 |
| GER_08 | 0.234 | 0.190 | 0.197 | 0.175 | 0.184 | 0.206 | 0.128 | 0.175 | 0.165 | 0.085 | 0.106 | 0.099 | 0.131 | 0.119 | 0.125 | 0.230 | 0.119 | 0.063 | 0.000 | | | | | GER_08 |
| GER_90 | 0.312 | 0.212 | 0.239 | 0.218 | 0.231 | 0.265 | 0.171 | 0.248 | 0.187 | 0.118 | 0.128 | 0.118 | 0.132 | 0.152 | 0.159 | 0.279 | 0.135 | 0.076 | 0.050 | 0.000 | | | | GER_90 |
| CAY_08 | 0.231 | 0.121 | 0.179 | 0.140 | 0.159 | 0.183 | 0.103 | 0.119 | 0.140 | 0.056 | 0.073 | 0.057 | 0.091 | 0.080 | 0.088 | 0.209 | 0.094 | 0.045 | 0.035 | 0.061 | 0.000 | | | CAY_08 |
| CAY_90 | 0.305 | 0.140 | 0.252 | 0.201 | 0.229 | 0.278 | 0.162 | 0.187 | 0.202 | 0.094 | 0.122 | 0.112 | 0.156 | 0.130 | 0.142 | 0.236 | 0.103 | 0.055 | 0.045 | 0.075 | 0.031 | 0.000 | | CAY_90 |
| LCA_08 | 0.246 | 0.164 | 0.195 | 0.165 | 0.186 | 0.239 | 0.127 | 0.158 | 0.153 | 0.073 | 0.101 | 0.081 | 0.113 | 0.096 | 0.105 | 0.209 | 0.107 | 0.053 | 0.025 | 0.044 | 0.029 | 0.030 | 0.000 | LCA_08 |

| | SLO_08 | PIS_08 | GRA_08 | MARI_08 | SHU_08 | ANT_08 | YNZ_08 | JAL_08 | COJ_06 | GAV_99 | REF_99 | TEC_02 | BUR_99 | MIS_02 | AND_02 | PAR_02 | CAR_99 | RIN_02 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SLO_08 | 0.000 | | | | | | | | | | | | | | | | | | SLO_08 |
| PIS_08 | 0.046 | 0.000 | | | | | | | | | | | | | | | | | PIS_08 |
| GRA_08 | 0.049 | 0.016 | 0.000 | | | | | | | | | | | | | | | | GRA_08 |
| MARI_08 | 0.108 | 0.063 | 0.050 | 0.000 | | | | | | | | | | | | | | | MARI_08 |
| SHU_08 | 0.081 | 0.132 | 0.136 | 0.137 | 0.000 | | | | | | | | | | | | | | SHU_08 |
| ANT_08 | 0.117 | 0.182 | 0.171 | 0.156 | 0.043 | 0.000 | | | | | | | | | | | | | ANT_08 |
| YNZ_08 | 0.161 | 0.189 | 0.202 | 0.275 | 0.157 | 0.310 | 0.000 | | | | | | | | | | | | YNZ_08 |
| JAL_08 | 0.195 | 0.219 | 0.251 | 0.322 | 0.186 | 0.342 | 0.045 | 0.000 | | | | | | | | | | | JAL_08 |
| COJ_06 | 0.317 | 0.306 | 0.333 | 0.395 | 0.288 | 0.426 | 0.169 | 0.135 | 0.000 | | | | | | | | | | COJ_06 |
| GAV_99 | 0.479 | 0.492 | 0.532 | 0.600 | 0.446 | 0.630 | 0.280 | 0.144 | 0.217 | 0.000 | | | | | | | | | GAV_99 |
| REF_99 | 0.643 | 0.643 | 0.736 | 0.801 | 0.595 | 0.791 | 0.418 | 0.270 | 0.367 | 0.140 | 0.000 | | | | | | | | REF_99 |
| TEC_02 | 0.515 | 0.513 | 0.583 | 0.646 | 0.484 | 0.665 | 0.325 | 0.198 | 0.252 | 0.071 | 0.063 | 0.000 | | | | | | | TEC_02 |
| BUR_99 | 0.537 | 0.530 | 0.570 | 0.669 | 0.527 | 0.725 | 0.318 | 0.201 | 0.240 | 0.074 | 0.212 | 0.115 | 0.000 | | | | | | BUR_99 |
| MIS_02 | 0.584 | 0.558 | 0.598 | 0.631 | 0.559 | 0.729 | 0.366 | 0.241 | 0.293 | 0.083 | 0.207 | 0.135 | 0.024 | 0.000 | | | | | MIS_02 |
| AND_02 | 0.650 | 0.616 | 0.644 | 0.672 | 0.595 | 0.763 | 0.426 | 0.291 | 0.388 | 0.124 | 0.243 | 0.207 | 0.087 | 0.048 | 0.000 | | | | AND_02 |
| PAR_02 | 0.408 | 0.391 | 0.421 | 0.469 | 0.404 | 0.537 | 0.274 | 0.167 | 0.185 | 0.056 | 0.220 | 0.140 | 0.066 | 0.061 | 0.104 | 0.000 | | | PAR_02 |
| CAR_99 | 0.611 | 0.598 | 0.642 | 0.716 | 0.553 | 0.734 | 0.387 | 0.227 | 0.275 | 0.037 | 0.160 | 0.099 | 0.084 | 0.091 | 0.133 | 0.058 | 0.000 | | CAR_99 |
| RIN_02 | 0.681 | 0.703 | 0.756 | 0.867 | 0.652 | 0.886 | 0.390 | 0.252 | 0.331 | 0.095 | 0.252 | 0.177 | 0.125 | 0.145 | 0.200 | 0.098 | 0.060 | 0.000 | RIN_02 |

# Appendix S2.9: Pairwise $F_{ST}$ for CCU (top) and COU (bottom).

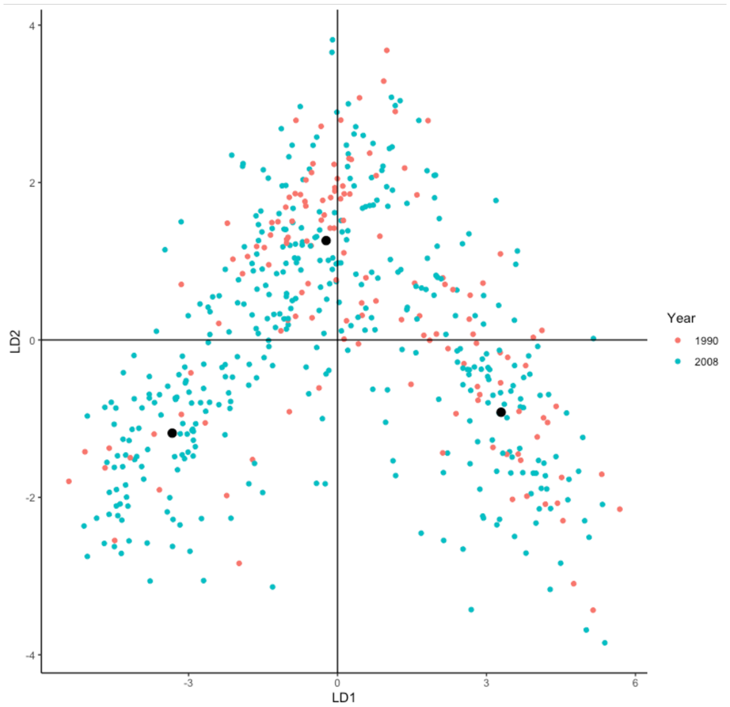
| | COR_08 | COR_90 | LAG_08 | LAG_90 | TOR_08 | TOR_90 | PUE_08 | PUE_90 | BRO_08 | LPC_08 | LPC_90 | PIC_08 | PIC_90 | SIM_08 | SIM_90 | ROS_08 | VIL_08 | VIL_90 | GER_08 | GER_90 | CAY_08 | CAY_90 | LCA_08 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| COR_08 | 0.000 | | | | | | | | | | | | | | | | | | | | | | | COR_08 |
| COR_90 | 0.122 | 0.000 | | | | | | | | | | | | | | | | | | | | | | COR_90 |
| LAG_08 | 0.040 | 0.087 | 0.000 | | | | | | | | | | | | | | | | | | | | | LAG_08 |
| LAG_90 | 0.060 | 0.077 | 0.026 | 0.000 | | | | | | | | | | | | | | | | | | | | LAG_90 |
| TOR_08 | 0.041 | 0.085 | 0.014 | 0.025 | 0.000 | | | | | | | | | | | | | | | | | | | TOR_08 |
| TOR_90 | 0.056 | 0.117 | 0.046 | 0.072 | 0.027 | 0.000 | | | | | | | | | | | | | | | | | | TOR_90 |
| PUE_08 | 0.054 | 0.074 | 0.024 | 0.030 | 0.027 | 0.046 | 0.000 | | | | | | | | | | | | | | | | | PUE_08 |
| PUE_90 | 0.144 | 0.100 | 0.090 | 0.084 | 0.091 | 0.116 | 0.051 | 0.000 | | | | | | | | | | | | | | | | PUE_90 |
| BRO_08 | 0.065 | 0.076 | 0.033 | 0.045 | 0.038 | 0.059 | 0.017 | 0.071 | 0.000 | | | | | | | | | | | | | | | BRO_08 |
| LPC_08 | 0.093 | 0.058 | 0.050 | 0.037 | 0.047 | 0.084 | 0.034 | 0.066 | 0.043 | 0.000 | | | | | | | | | | | | | | LPC_08 |
| LPC_90 | 0.110 | 0.063 | 0.063 | 0.052 | 0.055 | 0.095 | 0.041 | 0.079 | 0.052 | 0.015 | 0.000 | | | | | | | | | | | | | LPC_90 |
| PIC_08 | 0.109 | 0.063 | 0.058 | 0.033 | 0.048 | 0.093 | 0.041 | 0.074 | 0.055 | 0.017 | 0.028 | 0.000 | | | | | | | | | | | | PIC_08 |
| PIC_90 | 0.107 | 0.072 | 0.061 | 0.035 | 0.049 | 0.090 | 0.044 | 0.085 | 0.059 | 0.037 | 0.041 | 0.015 | 0.000 | | | | | | | | | | | PIC_90 |
| SIM_08 | 0.078 | 0.058 | 0.042 | 0.024 | 0.036 | 0.078 | 0.029 | 0.073 | 0.041 | 0.019 | 0.030 | 0.018 | 0.030 | 0.000 | | | | | | | | | | SIM_08 |
| SIM_90 | 0.084 | 0.061 | 0.046 | 0.030 | 0.039 | 0.073 | 0.034 | 0.071 | 0.044 | 0.021 | 0.036 | 0.021 | 0.032 | 0.011 | 0.000 | | | | | | | | | SIM_90 |
| ROS_08 | 0.124 | 0.087 | 0.090 | 0.082 | 0.094 | 0.142 | 0.096 | 0.123 | 0.094 | 0.069 | 0.093 | 0.085 | 0.108 | 0.062 | 0.060 | 0.000 | | | | | | | | ROS_08 |
| VIL_08 | 0.198 | 0.103 | 0.150 | 0.121 | 0.138 | 0.171 | 0.128 | 0.150 | 0.140 | 0.076 | 0.096 | 0.075 | 0.087 | 0.083 | 0.078 | 0.107 | 0.000 | | | | | | | VIL_08 |
| VIL_90 | 0.158 | 0.081 | 0.114 | 0.085 | 0.108 | 0.146 | 0.093 | 0.117 | 0.107 | 0.056 | 0.075 | 0.051 | 0.060 | 0.062 | 0.057 | 0.111 | 0.030 | 0.000 | | | | | | VIL_90 |
| GER_08 | 0.134 | 0.082 | 0.101 | 0.083 | 0.099 | 0.129 | 0.074 | 0.101 | 0.089 | 0.048 | 0.063 | 0.050 | 0.065 | 0.057 | 0.060 | 0.094 | 0.057 | 0.028 | 0.000 | | | | | GER_08 |
| GER_90 | 0.183 | 0.101 | 0.133 | 0.114 | 0.126 | 0.159 | 0.100 | 0.145 | 0.104 | 0.069 | 0.079 | 0.064 | 0.073 | 0.080 | 0.082 | 0.126 | 0.071 | 0.039 | 0.029 | 0.000 | | | | GER_90 |
| CAY_08 | 0.126 | 0.057 | 0.089 | 0.065 | 0.084 | 0.114 | 0.060 | 0.074 | 0.075 | 0.033 | 0.045 | 0.030 | 0.046 | 0.038 | 0.041 | 0.088 | 0.047 | 0.022 | 0.016 | 0.033 | 0.000 | | | CAY_08 |
| CAY_90 | 0.170 | 0.073 | 0.129 | 0.098 | 0.124 | 0.169 | 0.094 | 0.113 | 0.110 | 0.055 | 0.072 | 0.058 | 0.078 | 0.064 | 0.071 | 0.107 | 0.055 | 0.029 | 0.021 | 0.041 | 0.016 | 0.000 | | CAY_90 |
| LCA_08 | 0.140 | 0.074 | 0.102 | 0.079 | 0.101 | 0.143 | 0.075 | 0.096 | 0.084 | 0.042 | 0.060 | 0.042 | 0.056 | 0.049 | 0.053 | 0.093 | 0.051 | 0.024 | 0.012 | 0.025 | 0.013 | 0.015 | 0.000 | LCA_08 |

| | SLO_08 | PIS_08 | GRA_08 | MARI_08 | SHU_08 | ANT_08 | YNZ_08 | JAL_08 | COJ_06 | GAV_99 | REF_99 | TEC_02 | BUR_99 | MIS_02 | AND_02 | PAR_02 | CAR_99 | RIN_02 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SLO_08 | 0.000 | | | | | | | | | | | | | | | | | | SLO_08 |
| PIS_08 | 0.034 | 0.000 | | | | | | | | | | | | | | | | | PIS_08 |
| GRA_08 | 0.033 | 0.008 | 0.000 | | | | | | | | | | | | | | | | GRA_08 |
| MARI_08 | 0.067 | 0.030 | 0.022 | 0.000 | | | | | | | | | | | | | | | MARI_08 |
| SHU_08 | 0.049 | 0.067 | 0.069 | 0.079 | 0.000 | | | | | | | | | | | | | | SHU_08 |
| ANT_08 | 0.058 | 0.084 | 0.079 | 0.090 | 0.032 | 0.000 | | | | | | | | | | | | | ANT_08 |
| YNZ_08 | 0.106 | 0.109 | 0.116 | 0.144 | 0.090 | 0.169 | 0.000 | | | | | | | | | | | | YNZ_08 |
| JAL_08 | 0.131 | 0.129 | 0.143 | 0.171 | 0.111 | 0.179 | 0.034 | 0.000 | | | | | | | | | | | JAL_08 |
| COJ_06 | 0.199 | 0.180 | 0.191 | 0.210 | 0.168 | 0.225 | 0.100 | 0.098 | 0.000 | | | | | | | | | | COJ_06 |
| GAV_99 | 0.235 | 0.224 | 0.241 | 0.261 | 0.208 | 0.265 | 0.126 | 0.082 | 0.152 | 0.000 | | | | | | | | | GAV_99 |
| REF_99 | 0.321 | 0.294 | 0.321 | 0.337 | 0.280 | 0.332 | 0.193 | 0.164 | 0.247 | 0.089 | 0.000 | | | | | | | | REF_99 |
| TEC_02 | 0.237 | 0.221 | 0.243 | 0.262 | 0.207 | 0.262 | 0.138 | 0.103 | 0.157 | 0.043 | 0.057 | 0.000 | | | | | | | TEC_02 |
| BUR_99 | 0.245 | 0.227 | 0.244 | 0.269 | 0.223 | 0.280 | 0.133 | 0.099 | 0.165 | 0.044 | 0.127 | 0.066 | 0.000 | | | | | | BUR_99 |
| MIS_02 | 0.248 | 0.218 | 0.235 | 0.246 | 0.229 | 0.271 | 0.149 | 0.118 | 0.181 | 0.057 | 0.119 | 0.078 | 0.023 | 0.000 | | | | | MIS_02 |
| AND_02 | 0.256 | 0.224 | 0.239 | 0.249 | 0.230 | 0.271 | 0.166 | 0.135 | 0.206 | 0.075 | 0.134 | 0.101 | 0.058 | 0.030 | 0.000 | | | | AND_02 |
| PAR_02 | 0.191 | 0.170 | 0.184 | 0.200 | 0.178 | 0.220 | 0.120 | 0.089 | 0.121 | 0.040 | 0.139 | 0.078 | 0.040 | 0.035 | 0.059 | 0.000 | | | PAR_02 |
| CAR_99 | 0.272 | 0.252 | 0.270 | 0.289 | 0.239 | 0.291 | 0.158 | 0.112 | 0.188 | 0.025 | 0.098 | 0.059 | 0.052 | 0.059 | 0.079 | 0.043 | 0.000 | | CAR_99 |
| RIN_02 | 0.341 | 0.327 | 0.349 | 0.377 | 0.316 | 0.370 | 0.208 | 0.163 | 0.272 | 0.080 | 0.164 | 0.120 | 0.113 | 0.118 | 0.137 | 0.099 | 0.066 | 0.000 | RIN_02 |

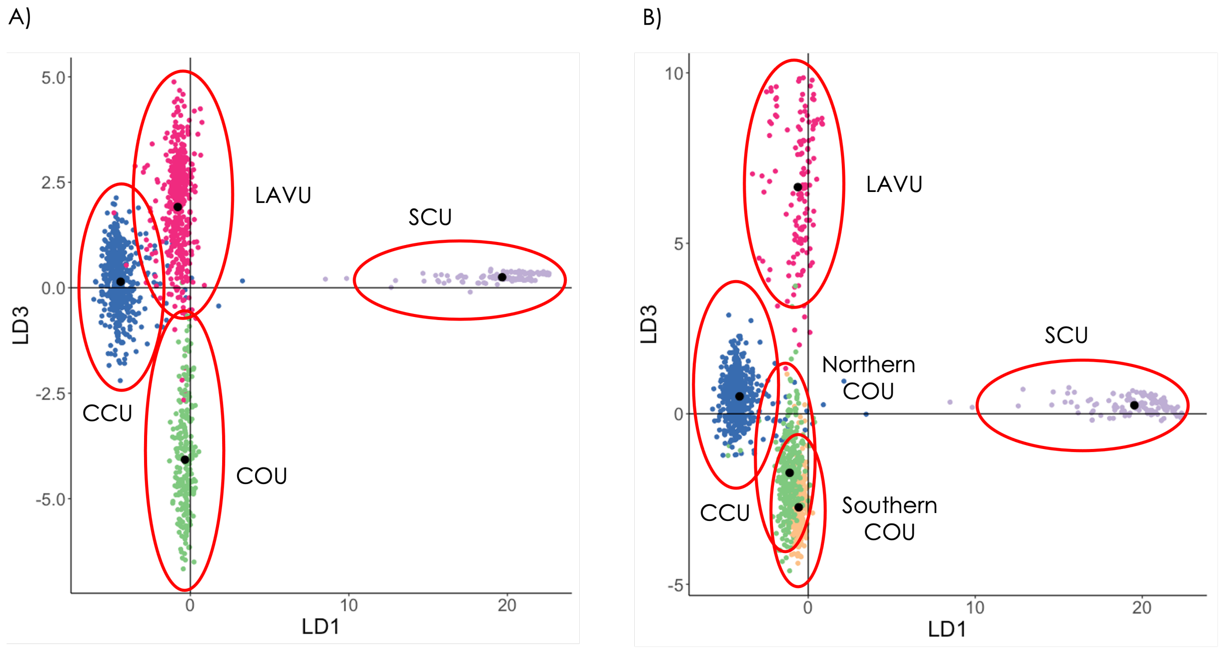
Appendix S2.10: Multiple linear regression (MLR) model output for dependent variable (Nei's unbiased genetic distance) and independent variables (Softer substrates, harder substrates, and number of headlands) for the COU dataset. Number of headlands has a significant positive correlation with Nei's genetic distance; however, this MLR output had a higher AIC compared to a separate MLR model using average kelp area as a third independent variable (see Appendix S2.4, Supporting Information). Asterisk represents the interaction between two or more variables. Outputs with significant p-values are in bold.

| Variable | Estimate | Std. Error | P-value | Eta squared ($\eta^2$) | Partial eta squared ($\eta^2_p$) | Omega squared ($\omega^2$) |
|---|---|---|---|---|---|---|
| Softer substrates | 2.0812e-6 | 5.908e-7 | 0.0006 | 0.02076 | 0.07883 | 0.01905 |
| **Harder substrates** | **9.8622e-6** | **1.874e-6** | **<0.0001** | **0.04634** | 0.16041 | 0.04459 |
| **Softer substrates * Harder substrates** | **4.557e-11** | **6.71e-11** | **0.4980** | **0.00077** | 0.00317 | 0.00001 |
| **Number of headlands** | **0.0327638** | **0.018183** | **0.0736** | **0.00543** | 0.02190 | 0.00375 |
| **Softer substrates * Number of headlands** | **-8.947e-8** | **3.613e-7** | **0.8048** | **0.00010** | 0.00042 | 0.00001 |
| **Harder substrates * Number of headlands** | **-3.388e-7** | **1.484e-6** | **0.8197** | **0.00009** | 0.00036 | 0.00001 |
| **Softer substrates * Harder substrates * Number of headlands** | **-4.31e-11** | **1.35e-11** | **0.0017** | **0.01708** | 0.06577 | 0.01538 |

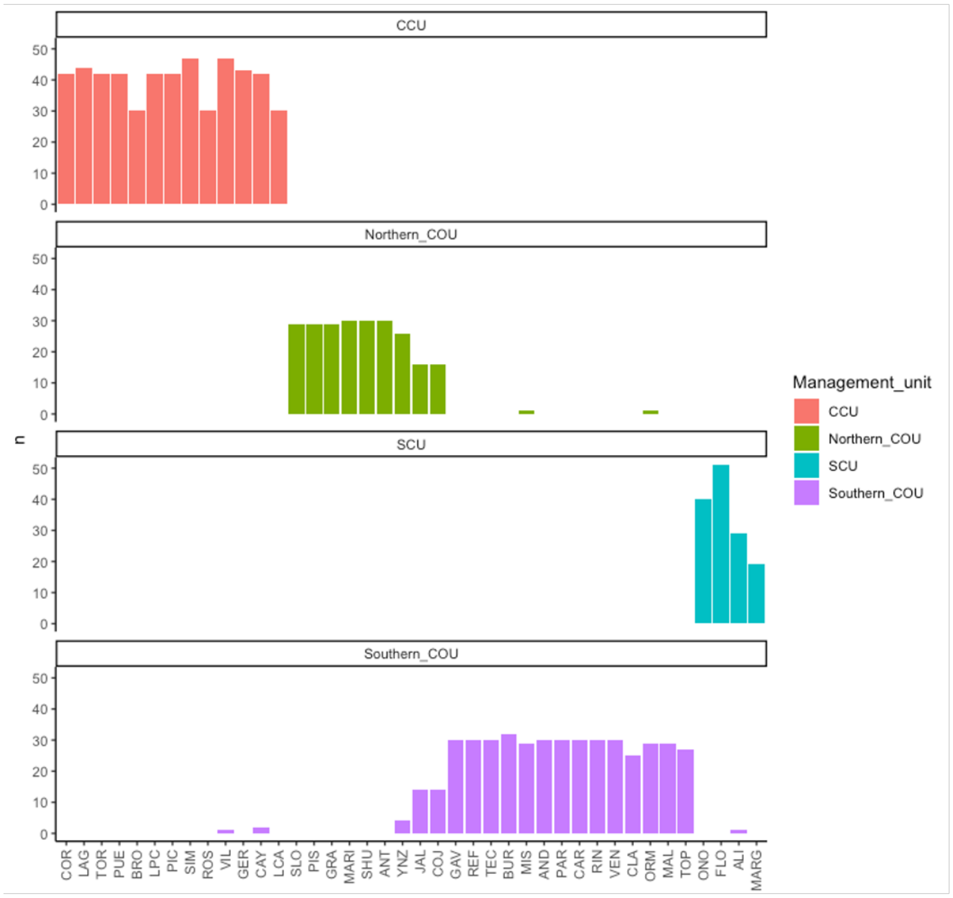
Appendix S2.11: DAPC (*K* = 3) for CCU showing no clustering between 1990 (red) and 2008 (blue) samples.



Appendix S2.12: Discriminant analysis of principal components (DAPC) for *K* = 4 (A) and *K* = 5 (B) for all four management units: CCU (blue), COU (green and orange), LAVU (rose), and SCU (purple). LD1 vs. LD3 are shown for each DAPC. (B) shows COU separated into two clusters. SCU is also separated from the CCU, COU, and LAVU clusters along the LD3 axis. This further corroborates that SCU supports a different species, *E. kristinae*, than the other three sites, which support *E. newberryi*.



Appendix S2.13: Distribution of individuals shows some samples had been clustered to a different management unit than where it was originally collected based on the DAPC (*K* = 4).



## References

Ahnelt, H., Göschl, J., Dawson, M. N., & Jacobs, D. K. (2004). Geographical variation in the cephalic lateral line canals of Eucyclogobius newberryi (Teleostei, Gobiidae) and its comparison with molecular phylogeography. *Folia Zoologica*, *53*(4), 358–398.

Amaral, A. R., Beheregaray, L. B., Bilgmann, K., Boutov, D., Freitas, L., Robertson, K. M., Sequeira, M., Stockin, K. A., Coelho, M. M., & Möller, L. M. (2012). Seascape Genetics of a Globally Distributed, Highly Mobile Marine Mammal: The Short-Beaked Common Dolphin (Genus Delphinus). *PLoS ONE*, *7*(2), e31482. https://doi.org/10.1371/journal.pone.0031482

Barlow, Michele. (2002). *Phylogeographic Structure of the Tidewater Goby, Eucyclogobius newberryi (Teleostei: Gobiidae), in the San Francisco Bay Area and Ventura County: Implications for Conservation Management* [Unpublished master's thesis]. University of California, Los Angeles.

Bell, T. W., Cavanaugh, K. C., Saccomanno, V. R., Cavanaugh, K. C., Houskeeper, H. F., Eddy, N., Schuetzenmeister, F., Rindlaub, N., & Gleason, M. (2023). Kelpwatch: A new visualization and analysis tool to explore kelp canopy dynamics reveals variable response to and recovery from marine heatwaves. *Plos One*, *18*(3), e0271477.

Bicknell, A. W. J., Knight, M. E., Bilton, D., Reid, J. B., Burke, T., & Votier, S. C. (2012). Population genetic structure and long-distance dispersal among seabird populations: Implications for colony persistence. *Molecular Ecology*, *21*(12), 2863–2876.

Blair, M. E., & Melnick, D. J. (2012). Scale-Dependent Effects of a Heterogeneous Landscape on Genetic Differentiation in the Central American Squirrel Monkey (Saimiri oerstedii). *PLoS ONE*, *7*(8), e43027. https://doi.org/10.1371/journal.pone.0043027

Charlesworth, B. (1998). Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution*, *15*(5), 538–543.

Cook, B. D., Kennard, M. J., Real, K., Pusey, B. J., & Hughes, J. M. (2011). Landscape genetic analysis of the tropical freshwater fish Mogurnda mogurnda (Eleotridae) in a monsoonal river basin: Importance of hydrographic factors and population history: Tropical river landscape genetic analyses. *Freshwater Biology*, *56*(5), 812–827. https://doi.org/10.1111/j.1365-2427.2010.02527.x

Dawson, M. N., Louie, K. D., Barlow, M., Jacobs, D. K., & Swift, C. C. (2002). Comparative phylogeography of sympatric sister species, Clevelandia ios and Eucyclogobius newberryi (Teleostei, Gobiidae), across the California Transition Zone. *Molecular Ecology*, *11*(6), 1065–1075.

Dawson, M. N., Staton, J. L., & Jacobs, D. K. (2001a). Phylogeography of the tidewater goby, Eucyclogobius newberryi (Teleostei, Gobiidae), in coastal California. *Evolution*, *55*(6), 1167–1179.

Dawson, M. N., Staton, J. L., & Jacobs, D. K. (2001b). Phylogeography of the Tidewater Goby, Eucyclogobius newberryi (Teleostei, Gobiidae), in Coastal California. *Evolution*, *55*(6), 1167–1179.

Earl, D. A., Louie, K. D., Bardeleben, C., Swift, C. C., & Jacobs, D. K. (2010). Rangewide microsatellite phylogeography of the endangered tidewater goby, Eucyclogobius newberryi (Teleostei: Gobiidae), a genetically subdivided coastal fish with limited marine dispersal. *Conservation Genetics*, *11*(1), 103–114.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method.

*Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Emaresi, G., Pellet, J., Dubey, S., Hirzel, A. H., & Fumagalli, L. (2011). Landscape genetics of the Alpine newt (Mesotriton alpestris) inferred from a strip-based approach. *Conservation Genetics*, *12*, 41–50.

Flores-Manzanero, A., Luna-Bárcenas, M. A., Dyer, R. J., & Vázquez-Domínguez, E. (2019). Functional connectivity and home range inferred at a microgeographic landscape genetics scale in a desert-dwelling rodent. *Ecology and Evolution*, *9*(1), 437–453.

Gu, Q.-H., Zhou, C.-J., Cheng, Q.-Q., Li, X.-J., Zhu, G.-R., Zhang, M., Gao, Y.-N., & Dong, J. (2015). The perplexing population genetic structure of Bellamya purificata (Gastropoda: Viviparidae): low genetic differentiation despite low dispersal ability. *Journal of Molluscan Studies*, *81*(4), 466–475.

Haugen, H., Linløkken, A., Østbye, K., & Heggenes, J. (2020). Landscape genetics of northern crested newt Triturus cristatus populations in a contrasting natural and human-impacted boreal forest. *Conservation Genetics*, *21*, 515–530.

Hellmair, M., & Kinziger, A. P. (2014). Increased extinction potential of insular fish populations with reduced life history variation and low genetic diversity. *PLoS One*, *9*(11), e113139.

Holderegger, R., Buehler, D., Gugerli, F., & Manel, S. (2010). Landscape genetics of plants. *Trends in Plant Science*, *15*(12), 675–683.

Homola, J. J., Loftin, C. S., & Kinnison, M. T. (2019). Landscape genetics reveals unique and shared effects of urbanization for two sympatric pool-breeding amphibians. *Ecology and Evolution*, *9*(20), 11799–11823. https://doi.org/10.1002/ece3.5685

Jenkins, D. G., Carey, M., Czerniewska, J., Fletcher, J., Hether, T., Jones, A., Knight, S., Knox, J., Long, T., Mannino, M., McGuire, M., Riffle, A., Segelsky, S., Shappell, L., Sterner, A., Strickler, T., & Tursi, R. (2010). A meta-analysis of isolation by distance: Relic or reference standard for landscape genetics? *Ecography*, no-no. https://doi.org/10.1111/j.1600-0587.2010.06285.x

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405.

Jombart, T., & Collins, C. (n.d.). *A tutorial for Discriminant Analysis of Principal Components (DAPC) using adegenet 2.0.0*.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 1–15.

Kalinowski, S. T. (2005). Do polymorphic loci require large sample sizes to estimate genetic distances? *Heredity*, *94*(1), 33–36.

Kinziger, A. P., Hellmair, M., McCraney, W. T., Jacobs, D. K., & Goldsmith, G. (2015). Temporal genetic analysis of the endangered tidewater goby: Extinction–colonization dynamics or drift in isolation? *Molecular Ecology*, *24*(22), 5544–5560.

Kunde, M. N., Martins, R. F., Premier, J., Fickel, J., & Förster, D. W. (2020). Population and landscape genetic analysis of the Malayan sun bear Helarctos malayanus. *Conservation Genetics*, *21*, 123–135.

Lafferty, K. D., Swenson, R. O., & Swift, C. C. (1996). Threatened fishes of the world: Eucyclogobius newberryi Girard, 1857 (Gobiidae). *Environmental Biology of Fishes*, *46*(3), 254–254. https://doi.org/10.1007/BF00005000

Lafferty, K. D., Swift, C. C., & Ambrose, R. F. (1999a). Extirpation and Recolonization in a Metapopulation of an Endangered Fish, the Tidewater Goby. *Conservation Biology*, *13*(6), 1447–1453. https://doi.org/10.1046/j.1523-1739.1999.98016.x

Lafferty, K. D., Swift, C. C., & Ambrose, R. F. (1999b). Extirpation and Recolonization in a Metapopulation of an Endangered Fish, the Tidewater Goby. *Conservation Biology*, *13*(6), 1447–1453. https://doi.org/10.1046/j.1523-1739.1999.98016.x

Lafferty, K. D., Swift, C. C., & Ambrose, R. F. (1999c). Postflood persistence and recolonization of endangered tidewater goby populations. *North American Journal of Fisheries Management*, *19*(2), 618–622.

Landguth, E. L., Cushman, S. A., Murphy, M. A., & Luikart, G. (2010). Relationships between migration rates and landscape resistance assessed using individual-based simulations. *Molecular Ecology Resources*, *10*(5), 854–862.

Leclerc, É., Mailhot, Y., Mingelbier, M., & Bernatchez, L. (2008). The landscape genetics of yellow perch ( *Perca flavescens* ) in a large fluvial ecosystem. *Molecular Ecology*, *17*(7), 1702–1717. https://doi.org/10.1111/j.1365-294X.2008.03710.x

Leon, C., Banks, S., Beck, N., & Heinsohn, R. (2022). Population genetic structure and dispersal patterns of a cooperative breeding bird in variable environmental conditions. *Animal Behaviour*, *183*, 127–137.

López-Márquez, V., Cushman, S. A., Templado, J., Wan, H. Y., Bothwell, H. M., & Machordom, A. (2021). Genetic connectivity of two marine gastropods in the Mediterranean Sea: Seascape genetics reveals species-specific oceanographic drivers of gene flow. *Molecular Ecology*, *30*(19), 4608–4629. https://doi.org/10.1111/mec.16080

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology & Evolution*, *28*(10), 614–621.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, *18*(4), 189–197.

Mapelli, F. J., Boston, E. S., Fameli, A., Gómez Fernández, M. J., Kittlein, M. J., & Mirol, P. M. (2020). Fragmenting fragments: Landscape genetics of a subterranean rodent (Mammalia, Ctenomyidae) living in a human-impacted wetland. *Landscape Ecology*, *35*, 1089–1106.

Martel, C. M., Sutter, M., Dorazio, R. M., & Kinziger, A. P. (2021). Using environmental DNA and occupancy modelling to estimate rangewide metapopulation dynamics. *Molecular Ecology*, *30*(13), 3340–3354. https://doi.org/10.1111/mec.15693

McCraney, W. T., Goldsmith, G., Jacobs, D. K., & Kinziger, A. P. (2010). Rampant drift in artificially fragmented populations of the endangered tidewater goby (Eucyclogobius newberryi). *Molecular Ecology*, *19*(16), 3315–3327.

Mendonca, H., Smith, J., & Brinegar, C. (2001). Isolation and Characterization of Four Microsatellite Loci in the Tidewater Goby (Eucyclogobius newberryi). *Marine Biotechnology*, *3*(2), 91–95. https://doi.org/10.1007/s101260000051

Milana, V., Franchini, P., Sola, L., Angiulli, E., & Rossi, A. R. (2012). Genetic structure in lagoons: The effects of habitat discontinuity and low dispersal ability on populations of Atherina boyeri. *Marine Biology*, *159*, 399–411.

Miller-Sims, V. C., Gerlach, G., Kingsford, M. J., & Atema, J. (2008). Dispersal in the spiny damselfish, Acanthochromis polyacanthus, a coral reef fish species without a larval pelagic stage. *Molecular Ecology*, *17*(23), 5036–5048.

Morinha, F., Dávila, J. A., Bastos, E., Cabral, J. A., Frías, Ó., González, J. L., Travassos, P., Carvalho, D., Milá, B., & Blanco, G. (2017). Extreme genetic structure in a social bird species despite high dispersal capacity. *Molecular Ecology*, *26*(10), 2812–2825.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, *70*(12), 3321–3323.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, *89*(3), 583–590.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—An update. *Bioinformatics*, *28*(19), 2537–2539.

Pritchard, J. K., Wen, W., & Falush, D. (2010). Documentation for STRUCTURE software: Version 2. *University of Chicago, Chicago, IL*.

R Core Team, R. (2013). *R: A language and environment for statistical computing*.

Ramos, R., Song, G., Navarro, J., Zhang, R., Symes, C. T., Forero, M. G., & Lei, F. (2016). Population genetic structure and long-distance dispersal of a recently expanding migratory bird. *Molecular Phylogenetics and Evolution*, *99*, 194–203.

Razgour, O., Rebelo, H., Puechmaille, S. J., Juste, J., Ibáñez, C., Kiefer, A., Burke, T., Dawson, D. A., & Jones, G. (2014). Scale-dependent effects of landscape variables on gene flow and population structure in bats. *Diversity and Distributions*, *20*(10), 1173–1185. https://doi.org/10.1111/ddi.12200

Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147.

Riginos, C., & Liggins, L. (2013). Seascape genetics: Populations, individuals, and genes marooned and adrift. *Geography Compass*, *7*(3), 197–216.

Rissler, L. J. (2016). Union of phylogeography and landscape genetics. *Proceedings of the National Academy of Sciences*, *113*(29), 8079–8086. https://doi.org/10.1073/pnas.1601073113

Segelbacher, G., Cushman, S. A., Epperson, B. K., Fortin, M.-J., Francois, O., Hardy, O. J., Holderegger, R., Taberlet, P., Waits, L. P., & Manel, S. (2010). Applications of landscape genetics in conservation biology: Concepts and challenges. *Conservation Genetics*, *11*, 375–385.

Selkoe, K. A., Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., von der Heyden, S., & Toonen, R. J. (2016). A decade of seascape genetics: Contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, *554*, 1–19.

Selkoe, K. A., Henzler, C. M., & Gaines, S. D. (2008). Seascape genetics and the spatial ecology of marine populations. *Fish and Fisheries*, *9*(4), 363–377.

Selkoe, K. A., Scribner, K. T., & Galindo, H. M. (2015). Waterscape Genetics—Applications of Landscape Genetics to Rivers, Lakes, and Seas. In N. Balkenhol, S. A. Cushman, A. T. Storfer, & L. P. Waits (Eds.), *Landscape Genetics* (pp. 220–246). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118525258.ch13

Sere, M., Thevenon, S., Belem, A. M. G., & De Meeus, T. (2017). Comparison of different genetic distances to test isolation by distance between populations. *Heredity*, *119*(2), 55–63.

Sork, V. L., & Smouse, P. E. (2006). Genetic analysis of landscape connectivity in tree populations. *Landscape Ecology*, *21*, 821–836.

Spies, B. T. (2014). *The effects of temperature and latitude on larval traits of the endangered tidewater goby (Eucyclogobius newberryi) and its sister species the arrow goby (Clevelandia ios)* [PhD Thesis]. California State University, Northridge.

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., Dezzani, R., Delmelle, E., Vierling, L., & Waits, L. P. (2007). Putting the 'landscape'in landscape genetics. *Heredity*, *98*(3), 128–142.

Storfer, A., Murphy, M. A., Spear, S. F., Holderegger, R., & Waits, L. P. (2010). Landscape genetics: Where are we now? *Molecular Ecology*, *19*(17), 3496–3514.

Swenson, R. O. (1997). Sex-role reversal in the tidewater goby, Eucyclogobius newberryi. *Environmental Biology of Fishes*, *50*, 27–40.

Swenson, R. O. (1999). The ecology, behavior, and conservation of the tidewater goby, Eucyclogobius newberryi. *Environmental Biology of Fishes*, *55*(1), 99–114.

Swift, C. C., Haglund, T. R., Ruiz, M., & Fisher, R. N. (1993). The status and distribution of the freshwater fishes of southern California. *Bulletin of the Southern California Academy of Sciences*, *92*(3), 101–167.

Swift, C. C., Nelson, J. L., Maslow, C., & Stein, T. (1989). *Biology and Distribution of the Tidewater Goby, Eucyclogobius Newberryi (Pisces: Gobiidae) of California*. Natural History Museum of Los Angeles County.

Swift, C. C., Spies, B., Ellingson, R. A., & Jacobs, D. K. (2016). A new species of the bay goby genus Eucyclogobius, endemic to southern California: Evolution, conservation, and decline. *PloS One*, *11*(7), e0158543.

U.S. Fish and Wildlife Services. (1994). *Endangered and Threatened Wildlife and Plants; Determination of Endangered Status for the Tidewater Goby*. https://www.govinfo.gov/content/pkg/FR-1994-02-04/html/94-2546.htm

U.S. Fish and Wildlife Services. (2005). *Recovery plan for the tidewater goby (Eucyclogobius newberryi)*. https://www.fws.gov/pacific/ecoservices/endangered/recovery/documents/TidewaterGobyfinalRecoveryPlan.pdf

U.S. Fish and Wildlife Services. (2013). *Endangered and Threatened Wildlife and Plants; Designation of Critical Habitat for Tidewater Goby*. https://www.govinfo.gov/content/pkg/FR-2013-02-06/pdf/2013-02057.pdf

U.S. Fish and Wildlife Services. (2014). *Endangered and Threatened Wildlife and Plants; Reclassifying the Tidewater Goby from Endangered to Threatened*. https://www.govinfo.gov/content/pkg/FR-2014-03-13/pdf/2014-05335.pdf

Waits, L. P., Cushman, S. A., & Spear, S. F. (2015). Applications of Landscape Genetics to Connectivity Research in Terrestrial Animals. In N. Balkenhol, S. A. Cushman, A. T. Storfer, & L. P. Waits (Eds.), *Landscape Genetics* (pp. 199–219). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118525258.ch12

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, *15*(1), 323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x

# Chapter 3

# Effects of drawing activities on undergraduate study behavior and midterm scores in a restoration ecology course

Benjamin A. Hà[1], Leryn Gorlitsky[1], Karen B. Givvin[2], Ji Y. Son[3], James W. Stigler[2]

[1] Department of Ecology and Evolutionary Biology, University of California, Los Angeles
[2] Department of Psychology, University of California, Los Angeles
[3] Department of Psychology, California State University, Los Angeles

## Abstract

Previous research has investigated the effects of drawing activities (DAs) on student learning in K-12 science courses, but less so in higher education. This research explores the effects of DAs in an ecological restoration course, both on learning outcomes and on whether students continue to use drawing after the intervention activity is over. Across two studies, students participated in two DAs over the span of a 10-week course. In the first study, pre- and post-survey responses (n=36) showed participants used their notes and lecture slides more than DAs to study for the midterm. In the second study, students exposed to the same drawing intervention were randomly instructed to either draw (drawers, n=23) or not draw (non-drawers, n=23) during a take-home midterm. Drawers were required to submit their drawings, but the quality of the drawings were not graded. Only the written responses from both groups were graded. Drawers scored significantly higher than non-drawers and had a more complex correlation network of paired words in their midterm responses. There was no correlation between the quality of the drawing and midterm scores within the experimental group. Together, these studies suggest DAs have the potential as an active learning tool to improve academic performance in undergraduate biology courses, though there is still much to learn about the best pedagogical practices for integrating DAs into a course.

## 1. Introduction

Visualizations are prevalent throughout science education and research. There have been various studies on how different types of visualizations impact students learning of science (Jenkinson, 2018; McElhaney et al., 2015; Stieff, 2017). For example, static images and diagrams can illustrate terminologies and concepts (e.g., components of an ecosystem, phases of a cell cycle, etc.), but have been found to limit students' understanding of biological processes compared to animations (McClean et al., 2005; Ryoo & Linn, 2012; Yarden & Yarden, 2010). Animations are a dynamic form of visualization that can be used as a tool to engage students in course content (Annetta et al., 2009; Korakakis et al., 2009). In research, scientists use graphical and textual visualizations to design experiments, represent data, and communicate results (Ainsworth et al., 2011; Kozma, 2003). However, instructional practices in the sciences, technology, engineering, and mathematics (STEM) have traditionally provided students with pre-generated visual-spatial representations in instructional materials (e.g., lectures, textbooks, etc.), which, in turn, reduces opportunities for students to construct their own visual representations of course content.

Learner-generated drawing is an active learning tool in which students can produce and discover their own interpretations of course material. We define drawing as a composite of freehand graphical and textual elements. The process of students drawing their own representations can facilitate problem solving, make missing information more explicit, represent implicit information explicitly, and set the foundation to prompt new understanding (Ainsworth & Loizou, 2003; Chi, 2009; Cox, 1999; Schwarz et al., 2009). Permitting students to generate their own representations of data can improve their understanding of representations (Ainsworth et al., 2011; diSessa, 2004), including learning the nature of scientific inquiry and how to communicate science (Chang et al., 2016; Prain & Tytler, 2013). However, some studies have also demonstrated that students do not always benefit from drawing (Ainsworth et al., 2016; De Bock et al., 2003; Leutner et al., 2009).

The impacts of drawing activities (DAs) on student learning have mostly been assessed in K-12 science classrooms (Dikmenli, 2010; Fiorella & Zhang, 2018; Gobert & Clement, 1999;

Hackling & Prain, 2005; Prain & Tytler, 2013; Quillin & Thomas, 2015; Van Meter & Garner, 2005) with fewer studies conducted in higher education biology courses (Cheng & Beal, 2020; Lin et al., 2017; Peart, 2022). Most relevant to our current study is one by Nugraha 2018, in which all students were required to draw in an upper-division undergraduate human physiology course. The author found a positive correlation between the quality of learner-generated drawings and written test scores; however, there was no control group for the treatment (Nugraha, 2018). Another relevant study is one by Heideman et al. 2017, which explored the study behavior of introductory biology undergraduates. Part of their experiments assessed if the participants adopted drawing as a study tool after using minute sketches with folded lists as a DA intervention; however, the authors used a different DA from our current study, and the experiments were designed in a non-classroom setting (Heideman et al., 2017).

Here, we report results from two separate studies conducted in an upper-division ecological restoration course at the University of California, Los Angeles (UCLA). The intervention consisted of a 10- to 12-minute DA where students drew in response to a provided prompt, and then engaged in small and large group discussions. The first study investigated students' tendencies to continue using drawing as part of their studying behavior after completing the DA intervention. The second study tested how drawing impacts student academic performance on a midterm using a randomized study design that compared an experimental group that engaged in a DA with a control group that did not. We hypothesized that there would be an increase in students using DAs as a study tool and that drawing would have a positive impact on academic performance.

## 2. Study 1

### *2.1 Methods*

#### *2.1.1 Course Description and Study Participants*

This study was conducted during an online lecture for EE BIOL 136: Ecological Restoration (hereafter, referred interchangeably as EE BIOL 136) in Spring 2020. The course was taught by second author, Dr. Leryn Gorlitsky, and had two discussion sections led by two different teaching assistants (UCLA graduate students). Each teaching assistant was responsible for grading assignments only for students enrolled in their respective discussion sections. Spring 2020 was the first pandemic quarter at UCLA where classes were taught remotely online.

Most students in this upper-division course were biology majors with the remainder majoring in marine biology, environmental science, or ecology, behavior, and evolution. 47 total undergraduates were enrolled and received extra credit for voluntarily participating in the research study. This research was approved by the UCLA Institutional Review Boards (#20-000410).

#### *2.1.2 Drawing Activity Intervention*

Two DAs were conducted during lecture to introduce the students to drawing as an active learning tool and a potential study tool. After receiving a lecture on assembly order, the students completed the first in-class DA during the subsequent lecture one week later. The second in-class DA followed one week after a separate lecture on ecological succession. Each DA had a prompt for students to draw an ecological restoration concept related to content from the previous lecture (Appendix S3.1, Supporting Information). For each DA, the guest instructor (i.e., the researcher), Benjamin Hà, first reviewed relevant content from the previous lecture, provided suggestions to students on how to visualize the content as a drawing, allocated about 10 minutes for students to draw the prompt, and then had students discuss their drawings in groups of 2-3. After the students discussed their drawings, the guest instructor presented a drawing they had made in advance under the same time constraints, and then led a whole-class discussion about what was right, wrong, and missing from the instructor's drawing. As part of the intervention, alternatives to drawing were recommended to the students to minimize their apprehension or unfamiliarity with drawing (see Appendix S3.1, Supporting Information). The same pedagogical structure was applied to each DA.

### *2.1.3 Midterm Format*

The midterm was a "take-home" and consisted of four open-ended (i.e., constructed-response) questions. Students were required to respond to the first two questions and then could choose to answer either the third or fourth question; students only had to respond to three out of the four questions. Each question challenged students to consider multiple restoration ecology concepts (e.g., assembly order, ecological succession, invasive species, etc.) to devise a restoration plan for managing or conserving an ecosystem.

### *2.1.4 Data Collection*

Participants anonymously completed pre- and post-surveys about their study behavior and past experiences in UCLA biology courses. Only 36 of the 47 enrolled students completed both the pre- and post-surveys (76.6% response rate). Only the responses from these participants were subsequently used for analysis.

The pre- and post-surveys had overlapping questions to compare differences or similarities in participant study behavior in previous UCLA biology courses and EE BIOL 136.

There were also questions about their past experiences, if any, with in-class DAs. Only the post-survey had questions about demographics. Survey questions included those on a 6-point Likert scale, short response, and open-ended.

### *2.1.5 Data Analysis*

Graph generation and statistical analyses (Wilcoxon rank sum test) were performed in R. The Wilcoxon rank sum test is a nonparametric alternative to the two-sample t-test. This was calculated to quantify variation in Likert survey responses regarding study behaviors from previous biology courses and in EE BIOL 136.

## 3. Results

### *3.1 Responses from pre- and post-surveys*

Based on anonymous self-reported responses from students who completed both the pre- and post-surveys, about 78.0% of all participants have participated in a DA in previous UCLA biology courses as part of a lecture whereas 22.0% have never participated in a DA before. About 25.6 % of all participants have never drawn as part of a response to an exam.

77.8% of participants claimed they studied prior to attempting the take-home midterm in Spring 2020, but the amount of time they spent studying is unknown. Students were asked to rank how often they used six study tools (personal notes, lecture slides, in-class drawing activities, reading assignments, online resources, and discussions with classmates) relative to one another in previous biology courses compared to Spring 2020 EE BIOL 136 (Figure 3.1). Students used their personal notes and the lecture slides the most, compared to the other four study tools, to learn the course material and to study for midterms. In-class DAs and assigned readings were used the least.

In general, students used their notes (Wilcoxon rank sum test, p-value < 0.0001) and the lecture slides (Wilcoxon rank sum test, p-value < 0.0001) significantly more than the in-class DAs, respectively, regardless of class. There was no statistically significant difference between using in-class DAs in previous biology courses compared to EE BIOL 136 to learn the course material (Wilcoxon rank sum test, p-value = 0.2388) or to study for midterms (Wilcoxon rank sum test, p-value = 0.8247). There was also no statistical difference within the other study tools to learn the course material or to study for midterms.

Despite the overall lack of use of in-class DAs outside of the classroom, 63.8% of the students reported that the exercise was helpful during class and 14.9% of the students drew other ecological restoration concepts on their own outside of class.

## 4. Discussion

### *4.1 Undergraduates use their notes and lecture slides the most compared to other study tools*

Our study found that students mostly reviewed their notes and the lecture slides to study for the midterm. This aligns with findings from Sebesta and Bray Speth 2017, which implemented self-regulated learning as an intervention where students set learning goals, monitor progress, and apply appropriate study strategies in an undergraduate introductory biology course (Sebesta & Bray Speth, 2017). Sebesta and Bray Speth 2017 found that students mostly reviewed their notes and lecture slides; however, reviewing notes and course materials (e.g., lecture slides) were grouped into the same category whereas our study separates them. In-class DAs were also not included as a strategy in Sebesta and Bray Speth 2017. On the other hand, Sebesta and Bray

Speth 2017 included other strategies, such as seeking assistance from peers, seeking assistance from other resources, and setting goals and tracking self-evaluation across the time of the intervention. While our study listed having discussions with classmates and using online resources as other potential studying strategies, the participants still used their notes and the lecture slides the most.

Our study also found that undergraduates are more likely to use their notes and lecture slides to study for midterms compared to in-class DAs. Respondents who anonymously self-reported that they studied prior to attempting the take-home midterm in Spring 2020 were significantly more likely to use both their notes and lecture slides than they were the in-class DAs in studying for the midterm. There was no significant difference in students reviewing their notes or the lecture slides in previous biology courses compared to EE BIOL 136, which suggests they have similar study behaviors across all biology courses.

While we originally hypothesized there would be an increase in students using in-class DAs after the drawing intervention, it is unsurprising that the students preferred to use their notes and lecture slides instead. Our findings contradict Heideman et al. 2017, which found a 10x increase in the use of drawing as a study tool after their intervention; however, Heideman et al. 2017 used a different DA than our study and was also conducted in a non-classroom setting. The authors implemented a DA called minute sketches with folded lists (see Heideman et al. 2017 for a full description of the activity). The main idea for this exercise was for users to use the minimum number of lines and symbols necessary to draw a concept in less than a minute. The paper is folded in such a way that users can only view one section containing one sketch at a time while hiding the other sections (Heideman et al., 2017; Quillin & Thomas, 2015).

Our study allotted 10 to 12 minutes for the students to draw based on a single drawing prompt, which permitted students to sketch one or more ecological concepts on one full piece of paper (8.5 x 11 inches). For each activity, the paper was folded in a way to allow students to draw more than one ecological concept (Appendix S3.1, Supporting Information). The activity was subsequently followed by think-pair-share (i.e., students discussed in groups of 2-3) and a large class discussion. We also conducted the study during the first quarter of classes being taught remotely online at UCLA due to the COVID-19 pandemic. Besides the difference in the DA intervention, there are many other potential reasons why our results contradicted those from Heideman et al. 2017.

One reason may be the variation in the environment for the participants. Heideman et al. 2017 conducted their study in a non-classroom setting whereas we conducted the study in the classroom. Variation in the environment may have different effects on student motivation. For example, those who volunteer to participate in an education research study in a non-classroom setting offers a low stakes environment where their behavior does not have a direct impact on their academic performance (i.e., grades). In contrast, participants who volunteer for a study conducted in a classroom setting may have more motivation to want to perform better since the stakes are higher.

Another reason may be the pedagogical approach. Heideman et al. 2017 incorporated practice sessions for the DA through a multi-day experiment; however, there does not seem to be a description of any pedagogy applied by an instructor during the intervention. In contrast, our study coupled the active learning tool with pedagogy led by an instructor, such as facilitating small and large group discussions. Although the variation in pedagogy may be a consequence of the research design. In other words, Heideman et al. 2017 may have been more focused on implementing a DA in a non-classroom setting whereas we wanted to conduct the intervention in a classroom, which would inevitably require an instructor. Regardless, it is important to consider the instructional support for enhancing learning through DAs (Van Meter & Garner, 2005; Wu & Rau, 2019).

### *4.2 Challenges to changing undergraduate study behavior in biology courses*

There is little research that tracks the extent to which students continue to use new study tools as part of their ongoing study routines after being introduced to an intervention. For example, Hensley et al. 2021 implemented an intervention that included a one-hour workshop focusing on effective learning strategies and self-awareness in the learning process (collectively referred to as the *Metacognition Treatment*) and another version of the intervention that also included a one-hour workshop on time management (referred to as *Metacognition+TM*) (Hensley et al., 2021)**.** Their study population was college students in an introductory biology course. Hensley et al. 2021 found an improvement in student exam scores and an increase in students using time management tools after the *Metacognition+TM* intervention.

In contrast, Sebesta and Bray Speth 2017 did not find a significant increase among undergraduates, who had scored D/F grades on two exams, in their self-evaluation (e.g., checking the progress of their work and understanding of the material) or in their setting goals and plans (e.g., making a timeline, keeping up with assigned reading or homework, etc.) (Sebesta & Bray Speth, 2017). However, this same group of students still had higher rates of seeking information (e.g., using outside resources not provided in class), seeking assistance from peers (e.g., studying with friends or classmates), and keeping records/monitoring (e.g., taking notes when studying from textbooks, lectures, etc.) – all of which align with more expected traditional approaches to studying compared to less common approaches (e.g., seeking instructor assistance, self-evaluation, goal setting, etc.).

These contrasting trends in students adopting different study strategies suggest a larger and more challenging research question: why or what precludes undergraduates from adopting non-traditional study strategies? When approaching this question, it would be important to consider external factors restricting students from adopting new study strategies (e.g., the general background of students, such as first-generation and/or lower socioeconomic status students who prioritize work or care of family members over studying, etc.) or the interactions of pedagogical structure with optimal study strategies. One possible study may be to identify whether or not participants change their study behavior if they are informed of the potential positive impacts of adopting a particular study behavior (e.g., improvement in grades, better engagement with course material, etc.). However, this might require introducing bias or deception into the study design, which would need to be carefully thought through.

## ***5. Study 2***

### *5.1 Overview*

Our second study investigates how drawing impacts student academic performance on a midterm using a randomized study design that compared an experimental group that was required to draw before providing a written response on a midterm compared to a control group that was not.

### *5.2 Methods*

#### *5.2.1. Course Description and Study Participants*

This project was conducted in EE BIOL 136: Ecological Restoration in Spring 2022. The course had a similar structure to the one in Study 1 (i.e., same instructor, same number of sections and teaching assistants, etc.), except Spring 2022 was the first quarter where classes returned to in-person teaching at UCLA.

The composition of student majors was similar to that in Study 1. In Spring 2022, 48 total undergraduates were enrolled and received participation points for voluntarily participating in the research study. This research was approved by the UCLA Institutional Review Boards (#20-000410).

### *5.2.2 Drawing Activity Intervention*

Two DAs were conducted in the course. They were similar in structure and pedagogy to the DA intervention implemented for Study 1 (Appendix S3.1, Supporting Information). The only difference was that the lectures on assembly order and ecological succession were merged. In turn, the first DA was conducted the week following the newly merged lecture and the second DA was two weeks following a lecture on invasive species. The guest instructor (i.e., the researcher), Benjamin Hà, was diagnosed with COVID-19 during this time, hence why the second DA was rescheduled two weeks following the respective lecture rather than one week as conducted in Study 1.

### 5*.2.3 Prompted Midterms*

Like the first study, the midterm was a "take-home" and comprised of constructed-response questions that covered multiple restoration ecology concepts. This time, though, the midterm had only two open-ended questions, and students were required to answer both questions. Midterms with varying instructions were randomly distributed to the students to establish an experimental and control group.

The experimental group was prompted to draw before constructing a written response, whereas the control group was not prompted to draw. The experimental group was also required to submit their drawings, but their drawing or drawing quality did not impact their score. Only the written responses for both groups were graded. The experimental group was also only prompted to draw for the first midterm question since students were given more freedom in their responses to the second question (i.e., students were allowed to propose a management plan for a toxic waste event at a location of their choice). Only some students (21.7%) chose to draw for the second question.

### *5.2.4 Data Collection*

Both midterm responses and drawings were collected for analysis. Two students were removed from the analysis: one student did not complete the midterm and the other student was an outlier in their midterm grade (Appendix S3.2, Supporting Information). This resulted in a final data set of 46 students for analysis (23 drawers and 23 non-drawers).

We did not collect data on the amount of time each student spent on their drawing due to the challenges of monitoring a take-home midterm. Two midterm drawings were lost during submission due to technical difficulties (i.e., Spring 2022 was also the first quarter UCLA transitioned to using a different online course website portal for submitting assignments). This resulted in analyzing 23 midterm responses and 21 drawings from the experimental group.

### *5.2.5 Data Analysis*

Graph generation, statistical analyses (Cohen's effect size and Pearson's correlation coefficient, Wilcoxon rank sum test, t-test), and qualitative analyses (word frequency, sentiment analysis, correlation network analysis) using the *tidytext* package (Silge & Robinson, 2016) were performed in R. Cohen's *d* indicates the effect size of differences between two groups by ignoring sample size and calculating the difference using standard deviations ($d = 0.2$ considered small effect size, 0.5 considered medium, and 0.8 considered large).

For data analysis in *tidytext*, midterm responses were formatted into one token per row, then filtered using stop words to remove common words and numbers. Word frequencies were calculated for each midterm. To normalize word frequencies, the total quantity of each word was divided by the sum of all words. This resulted in having one numerical value per word to represent its frequency (i.e., its proportion) relative to all words. We subsequently reviewed overlapping unique words found in the midterm responses for drawers and non-drawers, then identified which individual words were related to restoration ecology in a binary format.

To analyze sentiment, we used three lexicons that are well-known and have been used to test validity and reliability in qualitative data: 1) AFINN (Nielsen, 2011), which assigns words with a score ranging from -5 to 5 where negative scores indicate negative sentiment and positive scores indicate positive sentiment; 2) Bing et al. (Hu & Liu, 2004), which categorizes words in a binary format as negative or positive sentiments; and, 3) NRC Emotion (Mohammad & Turney, 2010, 2013), which categorizes words in a binary format across a spectrum of 10 different sentiments (positive, negative, anger, anticipation, disgust, fear, joy, sadness, surprise, and trust). One sentiment value per lexicon for each midterm response is the sum of the sentiment values using the aforementioned lexicon parameters.

A correlation network analysis was conducted to identify the correlation between paired words for the first midterm question. This shows which two words were more likely to appear together. This analysis was conducted separately for the first midterm question as well as across the entire midterm.

## ***6. Results***

### *6.1 Impact of drawing on midterm scores*

The maximum number of points for the midterm was 80, 40 for each question. The class average for the midterm was 89.6%. There was no statistically significant difference in the Spring 2022 EE BIOL 136 midterm scores between the two discussion sections (t-test, p-value = 0.9122; Figure 3.2a); however, there was a statistically significant difference between the drawers and non-drawers (t-test, p-value < 0.05; Figure 3.2b) with a moderate effect size (Cohen's effect size, $d = 0.4825$). The average midterm score was 91.4% for drawers and 87.8% for non-drawers.

### *6.2 Qualitative analysis of midterm responses and drawings*

Qualitative analyses included assessing word frequencies and conducting a sentiment and correlation network analysis on all midterm responses. Overall, drawers used more words (n = 24,693) than non-drawers (n = 23,085), but there was no statistical difference between the two groups (t-test, p-value = 0.5522). Initially, the words drawers and non-drawers used in their midterm responses were highly correlated (Appendix S3.3, Supporting Information) for the first (Pearson's correlation coefficient, $r = 0.9685$) and second midterm questions (Pearson's correlation coefficient, $r = 0.9181$). Subsequently, after categorizing each unique word found in common between drawers and non-drawers that was related to restoration ecology (n = 954), drawers used significantly more of these words than non-drawers (t-test, p-value < 0.005). Figure 3.3 shows the top 20 most common restoration ecology words between drawers and non-drawers. Variations of a word within the same part of speech were combined (e.g., different conjugations of the same verb, singular vs. plural nouns, etc.).

Sentiment analysis revealed variation in negative and positive sentiment values for AFINN and NRC Emotion lexicons for both groups. There was an overall negative average of sentiment values for the Bing et al. lexicon for both groups (Appendix S3.4, Supporting Information). Drawers had a higher range of variation in sentiment values compared to non-drawers, but there was no statistical significance between drawers and non-drawers for AFINN (t-test, p-value = 0.4884), NRC Emotion (t-test, p-value = 0.6074), nor Bing et al. (t-test, p-value = 0.9843) lexicons. There was also no correlation between the level of drawing quality and midterm scores (Pearson's correlation coefficient, $r = 0.19$) (Appendix S3.5, Supporting Information).

A correlation network analysis demonstrated that drawers had a more complex and higher range in correlation between paired words than non-drawers for the first midterm question (Figure 3.4). This analysis focuses on the first midterm question since it was the only question that required drawing; however, similar results are found across the entire midterm (Appendix S3.6, Supporting Information).

## *7. Discussion*

### *7.1 Drawing may function similarly to outlining before writing*

Based on our results, drawers scored significantly higher than non-drawers on the midterm overall. This suggests that drawing has a positive effect on midterm scores, with a moderate effect size. The lack of variation in sentiment values between the drawer and non-drawer midterm responses does not seem to provide insight into the difference in midterm scores. Furthermore, the quality of the drawings does not seem to explain the variation in midterm scores among the drawers. Our findings contradict Nugraha 2018, in which the author found a positive correlation between drawing quality and test scores (Nugraha, 2018). Nugraha 2018 had no control group and only allotted 15 minutes total for drawing to occur. Additionally, we assume the drawing was conducted before the written test, but this detail remains unclear. In the current study, we included a control group but did not instruct the experimental group to draw for a minimum amount of time due to the challenges of monitoring a take-home midterm. Our study also had a smaller sample size given that the drawing assessment only included the experimental group. As a result, these factors may have affected the quality of the drawings and in turn, its relation to midterm scores.

One reason why drawers performed better than non-drawers on the midterm may be related to the results demonstrating that drawers had a more complex correlation network and a higher correlation between paired words compared to non-drawers. This suggests drawers used words more interconnectedly whereas non-drawers used words more linearly. This appears to be true for the first midterm question and across the entire midterm.

Another reason why drawers performed better may be that the act of drawing before writing aided drawers in reducing cognitive load and organizing their thought process before constructing their responses. This further suggests drawing may function similarly to outlining before writing given that outlining has been shown to improve undergraduate writing quality (Kellogg, 1988; Torrance et al., 2000).

Outlining can reduce cognitive load while writing a text (De Smet et al., 2011; Kellogg, 2008). Cognitive load refers to the amount of mental effort a learner expends, and is a function of the number of novel elements that need to be kept in working (i.e., short-term) memory and the degree of interaction between those novel elements (Kirschner, 2002). For example, outlining has been observed to produce higher quality essays in psychology college students (Kellogg, 1988). Additionally, Torrance et al. 2000 found that psychology undergraduates who had a "minimal-drafting" approach to writing essays (i.e., producing one or at most two drafts) had lower quality essays compared to students who performed "outline-and-develop" (i.e., developed content prior to and during drafting) and "detailed-planning" strategies (i.e., using multiple content-development approaches, including outlining, mind mapping, brainstorming, rough drafting, etc.). This implies that having some sort of structured approach prior to writing essays can improve writing quality. To our knowledge, there have been no studies investigating the effects of outlining in relation to academic performance on exams with constructed-response questions in biology higher education courses, or in STEM in general.

### *7.2 Moving forward with improving undergraduate writing*

If outlining before writing can improve writing quality, then why are we, as instructors, creating constructed-response questions as part of exams without offering students the opportunity or time to produce higher-quality written responses during an exam? Part of this dilemma may be related to instructors failing to teach undergraduates how to write or improve their scientific writing, or perhaps there may be an institutional lack of courses or resources focused on scientific writing (Jerde & Taper, 2004). Another part of the issue may be instructors who focus on instructor-centered pedagogy (i.e., the instructor judges what information is legitimate or important and defines what constitutes as knowledge and learning) rather than student-centered pedagogy (i.e., the construction of knowledge is shared and learned through student engagement and evidence-based teaching approaches). This permits instructor-centered teaching approaches to prioritize grades over the learning process, which is a bit ironic in that student-centered pedagogies can improve grades (Armbruster et al., 2009; Connell et al., 2016; Mostrom & Blumberg, 2012). The challenge of grading students based on their scientific writing becomes a greater issue as science faculty continue to integrate constructed-response questions in "some" or "all" of their classes (Goubeaud, 2010) while simultaneously not educating undergraduates on how to produce scientific writing.

Biology instructors employ a spectrum of writing exercises in their courses (e.g., lab reports, summaries, research papers, etc.), yet we do not integrate instruction on how to write scientifically or what constitutes as good writing into the curriculum. Another related challenge may be instructors informing students about campus resources (i.e., writing centers), but the students choosing to not access these resources due to personal or logistical conflicts. The fundamental process of and skills for scientific research entail scientific writing and communication, in addition to numerous other skills (e.g., critical analysis of literature, data interpretation, problem solving, experimental design, etc.). While a study reveals that faculty support teaching undergraduates the skills related to the scientific process, they typically do not spend enough time teaching these skills due to the perception of needing to prioritize course content (Coil et al., 2010).

One way this can be appeased is for instructors to develop inquiry-based courses coupled with student-centered pedagogy in higher education, which grants instructors the flexibility to develop innovative ways of engaging and teaching undergraduates, especially regarding scientific writing (Justice et al., 2006; Newton & Tonelli, 2020). For undergraduate biology courses, some studies have found improvement in scientific writing when: 1) providing students with basic information on writing style and options for organizing and writing research papers (Justice et al., 2006); 2) including peer assessment as part of iterations for a biology report assignment (Liang & Tsai, 2010); and, 3) teaching about the peer-review process for scientific publications by including a collaborative review paper as a course assignment to be submitted for peer-review (Guilford, 2001). Another approach could also be to show students examples of good vs. bad published writing and review why that may be the case for each example. There is a need to educate students on how to produce high-quality scientific writing, and part of this process should entail teaching students the benefits of outlining before writing (De Smet et al., 2011; Kellogg, 1988, 2008; Torrance et al., 2000).

Investigating the effects of outlining on writing quality for constructed-response questions on biology exams in higher education may be an avenue for future research. Instructors should also reflect on why they want to integrate constructed-response questions into their exams. This latter point is especially important given that there may be differences in grades for constructed-response exams based on gender (Federer et al., 2016) and socioeconomic status (SES) where middle/high-SES students were favored as the proportion of constructed-response questions on exams increased for undergraduates in an introductory biology course (Wright et al., 2016).

### *8. Limitations and Future Directions*

One major caveat to our study is the take-home midterm, which allowed students more time beyond the scheduled class time to complete the exam. Our study also does not account for the total amount of time drawers spent drawing before writing. It is possible our results would vary for in-person exams or if we were to allot the time for drawing to occur separate from the total time allotted to complete the exam itself. An improvement to this study may be to conduct an in-person exam with constructed-response questions, grant the experimental group a fixed amount of time (~10 minutes) to only draw before writing, and then still allow the experimental group to complete the midterm under the same time constraints as the control group. Although, then the instructor would need to address the logistical challenges in scheduling exams between the two groups. It would also be beneficial to have conducted the study using a larger sample size, which could improve the statistical significance in grades between the two groups. Additionally, having a greater difference in the mean and standard deviation of the midterm scores between the two groups has the potential to improve the effect size.

Another caveat to the study may be the variation in prior knowledge of restoration ecology concepts across the participants. While we did not measure prior knowledge, EE BIOL 136 is an upper-division course in the Department of Ecology and Evolutionary Biology at UCLA. Any general concepts regarding restoration ecology may have either been covered in lower-division pre-requisite courses or in other upper-division courses or electives, in which case would affect students' prior knowledge and, in turn, their academic performance. For example, there have been many studies that found evidence for prior knowledge positively impacting academic achievement in higher education (Hailikari et al., 2008; Martin et al., 2013; Thompson & Zamboanga, 2004).

Otherwise, we are not too concerned about the variation in the study participants' confidence in drawing. Only 21.4% of the students reported that they have never participated in an in-class DA in previous biology courses. To account for this, the in-class DA for both studies included guidance on how to draw to reduce students' apprehension with drawing. While we are not too concerned with the students' confidence in drawing, there may be variation in any given students' drawing ability to translate written text into a visual-spatial representation. For example, the quality of learning for the remaining students who did have prior experience with in-class DAs in other biology courses remains unclear. In-class DAs are useful since they grant students the opportunity to create their own representations of course content while enabling instructors to accommodate a wide range of topics. Using Bloom's hierarchical model for learning as a reference, in-class DAs can cover lower-order skills, such as labeling terminology (e.g., muscles in human anatomy, segments of an invertebrate, etc.), and higher-order skills, such as drawing two or more biological concepts (e.g., showing how the internal process of photosynthesis in a plant relates to the carbon cycle in an ecosystem, etc.). For the first study, we did not investigate the details of any given students' level of experience with DAs in previous biology courses, but some students reported participating in DAs by labeling structures in plants and/or invertebrates. Students who did have prior experience with in-class DAs may have been better prepared to translate written prompts into visual-spatial representations, although the randomization of the midterms would have, in theory, accounted for this potential bias.

While there is growing empirical support for learning by drawing (Fiorella & Mayer, 2016; Fiorella & Zhang, 2018; Schmeck et al., 2014; Van Meter et al., 2006), not all studies have reported positive effects (Ainsworth et al., 2016; De Bock et al., 2003; Leutner et al., 2009). It is important to consider the specific type of DA being implemented as well as variation in the instructor’s pedagogy to supplement learning with DAs (Van Meter & Garner, 2005; Wu & Rau, 2019). It is also important for studies to have an appropriate research design to compare experimental and control groups, which can yield different results (Fiorella & Zhang, 2018).

In summary, our study expands on education research in higher education biology courses. The two studies reported here build on the literature exploring the effectiveness of DAs on academic performance and the presence, or absence, of impact an intervention has on modifying undergraduate study behavior. Rotating in-class DAs as part of an instructor’s repertoire to implement active learning tools can be refreshing for the instructor and students. Continuing to integrate active learning can help improve student academic performance (Freeman, 2014; Freeman et al., 2007), especially when coupled with student-centered pedagogy (Armbruster et al., 2009; Connell et al., 2016; Ebert-May et al., 1997; Freeman et al., 2007; Knight & Wood, 2005; Udovic et al., 2002; Walker et al., 2008). It is imperative for instructors to strive to continue expanding their pedagogical development and employing active and inclusive teaching strategies if we are to work collectively toward improving student learning.

Figure 3.1: Students ranked the six study tools on a 6-point Likert scale from most used (6) to least used (1) in previous biology courses (beige) and in the Spring 2020 EE BIOL 136: Ecological Restoration course at UCLA (coral). The study tools are shown in descending order. Note the change in the order of study tools when students used them to study for the midterm (A) and to learn course material (B). Overall, students used their notes (Wilcoxon rank sum test, p-value < 0.0001) and the lecture slides (Wilcoxon rank sum test, p-value < 0.0001) significantly more than the in-class DAs, respectively, regardless of the class. There was no statistically significant difference between using in-class DAs in previous biology courses compared to EE BIOL 136 to learn the course material (Wilcoxon rank sum test, p-value = 0.2388) or to study for midterms (Wilcoxon rank sum test, p-value = 0.8247).

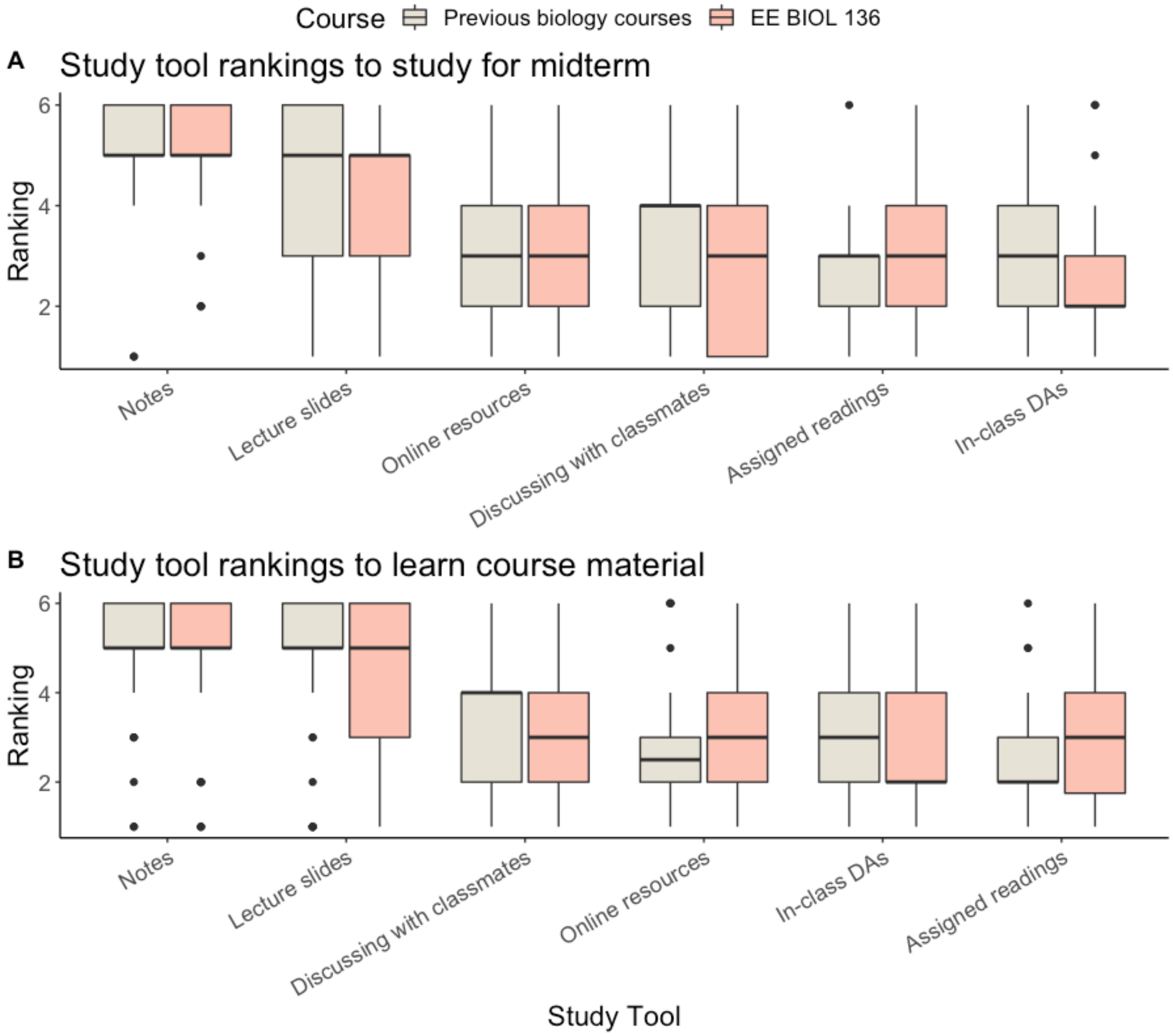


Figure 3.2: Midterm scores for students by discussion section (A) and treatment vs. control groups (B). Red dotted line denotes class average (89.6%). There was no statistical difference in grades between discussion sections (t-test, p-value = 0.9122), but there was a statistical difference in grades between groups (t-test, p-value < 0.05).

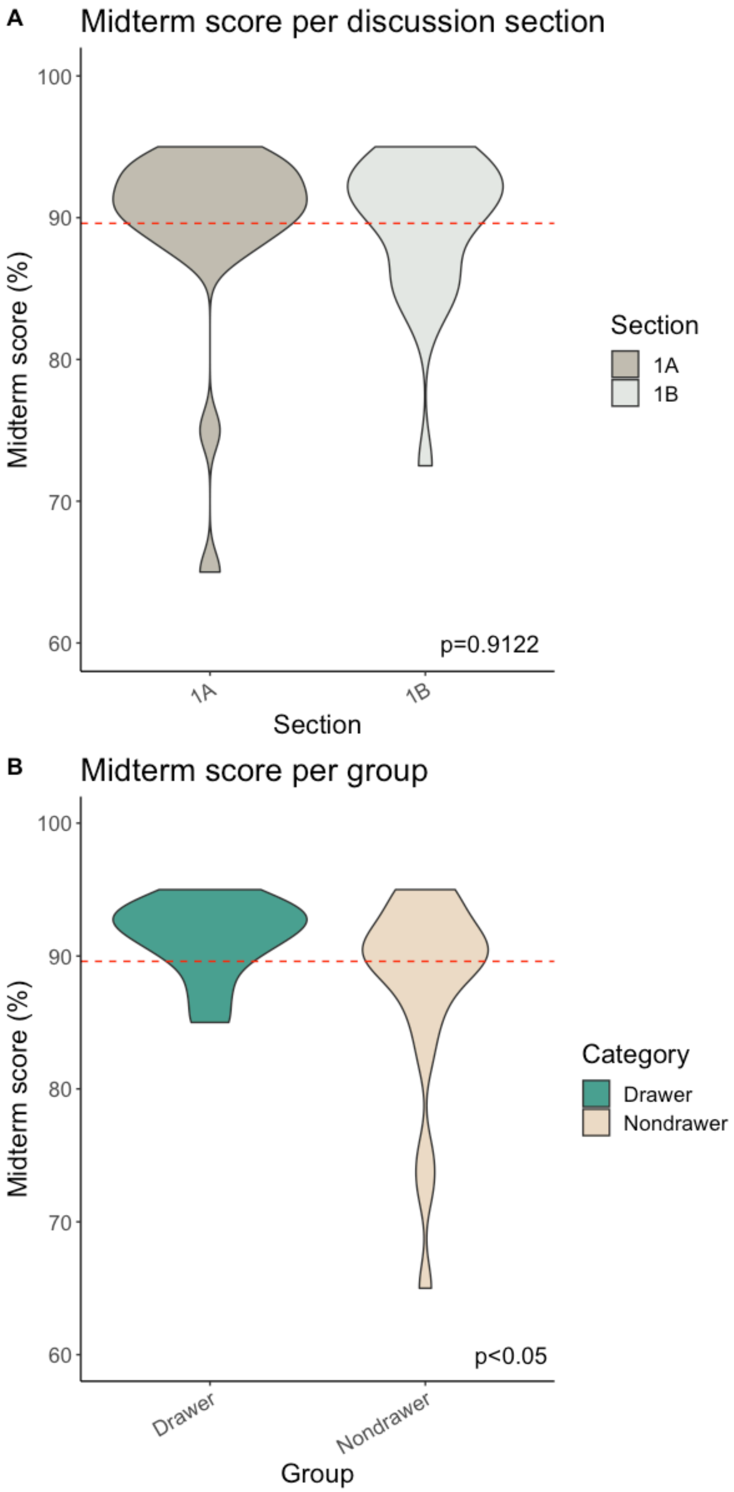


Figure 3.3: Drawers used significantly more words related to restoration ecology than non-drawers (t-test, p-value < 0.005). Asterisk denotes when multiple variations of a word had been combined (i.e., different conjugations of the same verb, singular vs. plural nouns, etc. were all consolidated within their respective types).

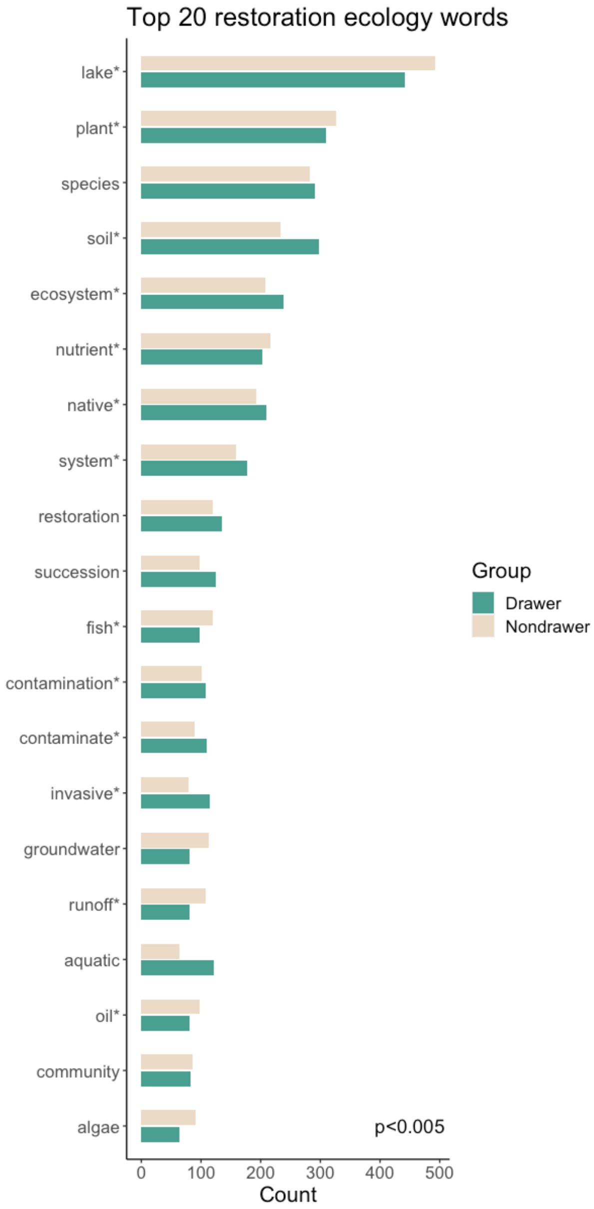
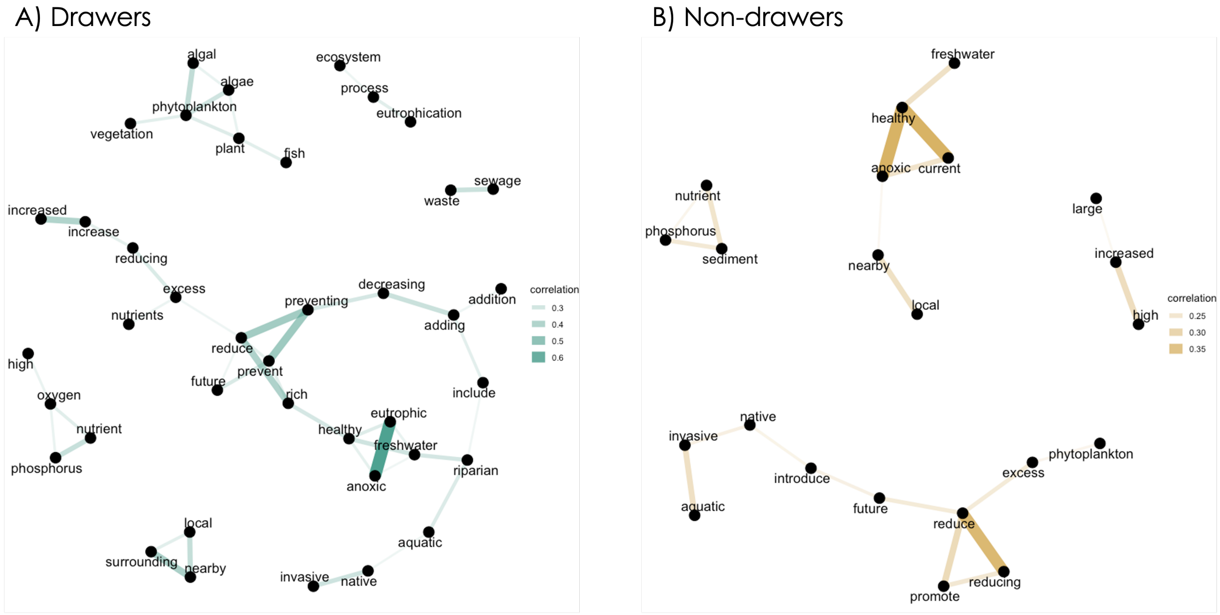


Figure 3.4: Correlation network analysis for paired words for the first midterm question response between drawers (A) and non-drawers (B). Drawers have a more complex and higher correlation between paired words for the first midterm question response compared to non-drawers. Note the difference in numeric scale within each group.



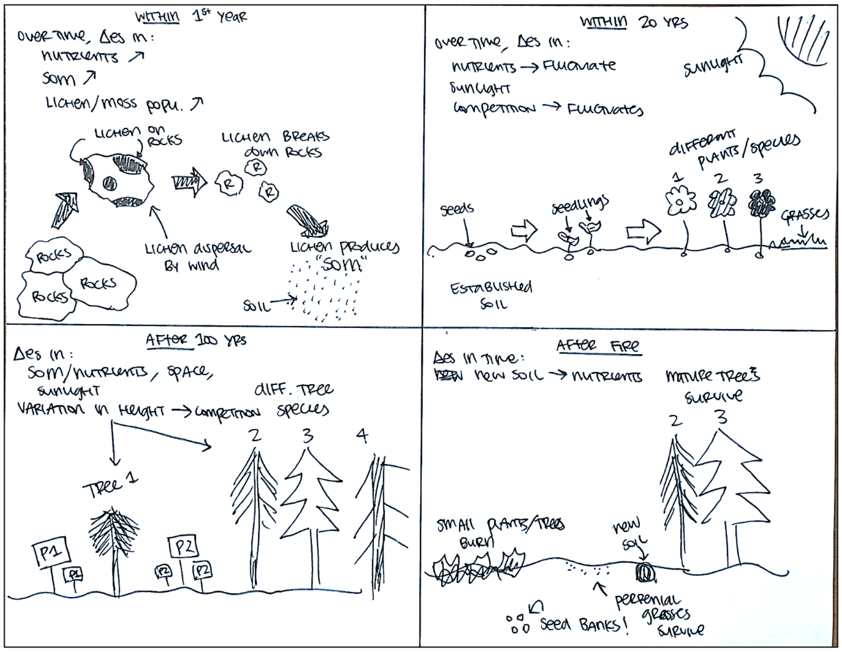
A) Drawers

B) Non-drawers

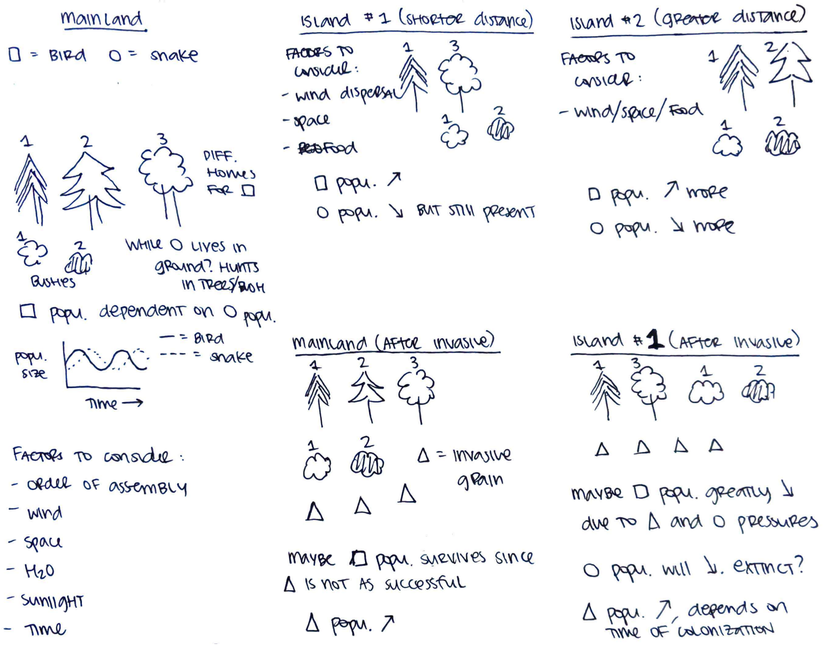
Appendix S3.1: Description of the drawing prompts for the two in-class drawing activities and the pedagogical approaches applied.

- Drawing prompt #1 (Restoration Ecology Theme: Succession):
  - Fold your paper so that you have 4 equal subsections
  - You will have ~6-8 minutes to draw the prompt (~1-2 min per ¼ subsection)
  - Using one writing utensil, draw your first interpretation of the following:
    - A volcano nearby a coastline recently erupted from the depths of the ocean and created a new island. In each ¼ section of your paper, draw what could happen in the area in the following situations:
      - Area <u>within</u> the first year. Hint: What's happening in the soil?
      - Area <u>within</u> 20 years. Hint: Think about plant seedlings
      - Area <u>after</u> 100 years. Hint: Think about other species
      - <u>After</u> a frequent/annual fire. Hint: What's happening post-fire?
- Drawing prompt #2 (Restoration Ecology Theme: Assembly Order and Invasive Species):
  - Fold your paper into thirds so that you have 3 columns
  - You will have ~10-12 minutes to draw the prompt
  - Using one writing utensil, draw your first interpretation of the following:
    - Part A (~5-7 minutes): There is one mainland and two islands at different distances from the mainland. The mainland is home to a granivorous bird and snake species. Given that the bird's eggs are prey for the snake, draw what you think might happen to each population
    - Part B (~5 minutes): Draw what happens when an invasive grain (from a grass) colonizes the mainland and the closest island
    - Hint: Consider how the rates for each species may change over time
- Pedagogy for leading in-class drawing activities:
  - Review lecture (5 minutes)
  - Suggested alternatives to drawing (5 minutes)
    - If you cannot "draw" something, you can box words, use arrows, draw simple line graphs, or use symbols to represent anything. For example, you can draw shapes to represent animals (e.g., a square denotes a bird, and a circle denotes a plant, etc.), use arrows to show directionality (e.g., population increases or decreases in size, etc.), or use numbers to represent different species for an organism (e.g., plant 1 vs. plant 2, etc.)
  - Students draw prompt individually (about 10 minutes depending on the prompt)
  - Think-pair-share (10 minutes)
  - Share the photo of the instructor's attempt at the in-class drawing activity under the same time constraints (see below) and lead a large class discussion (30-50 minutes)
    - Analyzed the instructor's drawing to discuss: what is right, what is wrong/missing, what is misleading about the prompt (e.g., what are your assumptions)
    - Pedagogy applied: Instructor guided students based on inquiry-based instruction to demonstrate how multiple restoration ecology concepts were related to the prompt

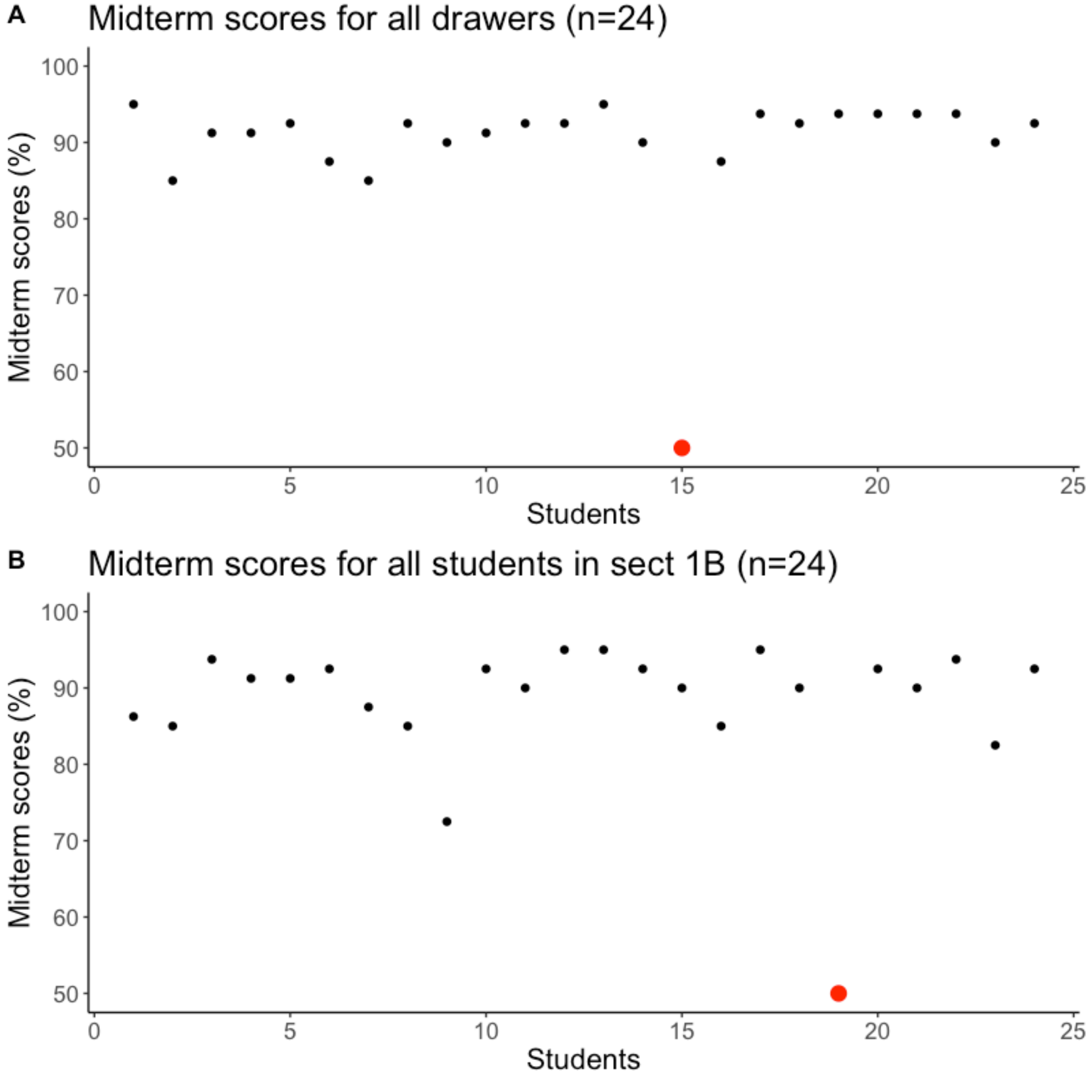
- Instructor’s attempting at drawing prompts:
  - Succession:



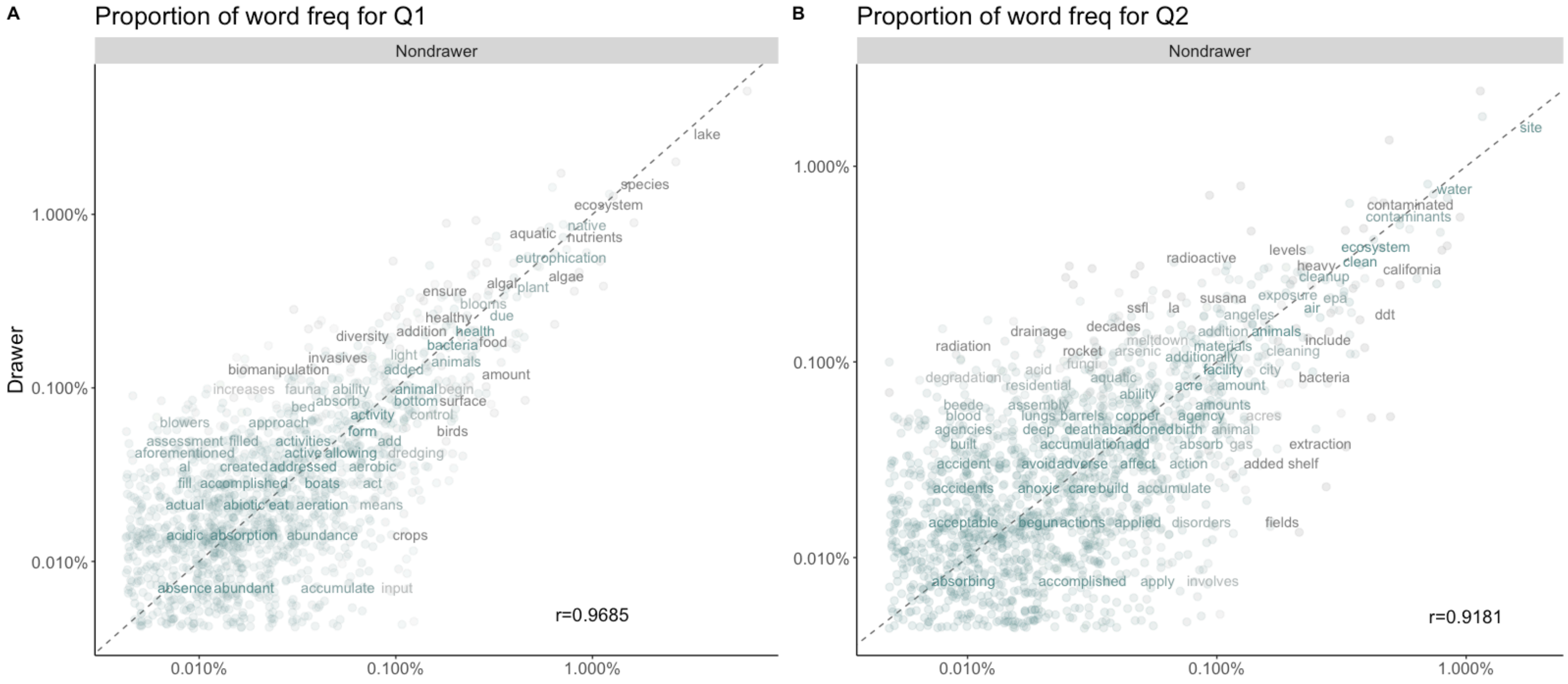
  - Assembly Order and Invasive Species:



Appendix S3.2: Two students were removed from the study. The first student (not shown) had an incomplete midterm. Red denotes the second student as an outlier for the midterm score in their discussion section (A) and the experimental group (B). For confidentiality, each student's name has been replaced by a participant number.

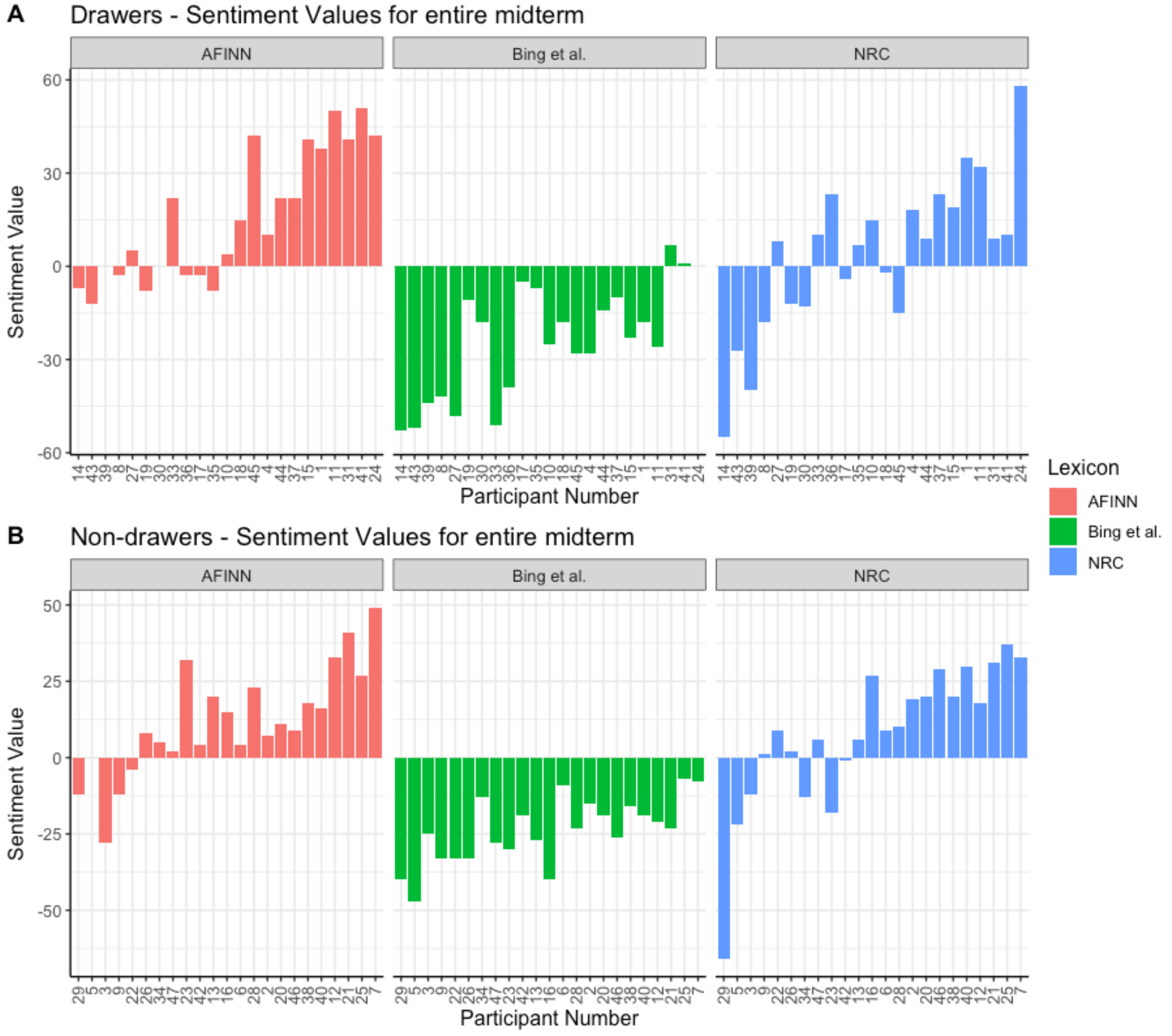


Appendix S3.3: After removing common words (i.e., stop words) and numbers, word frequencies were calculated for all midterm responses. To normalize word frequencies, the total quantity of each word was divided by the sum of all words. This resulted in having one numerical value per word to represent its frequency (i.e., its proportion) relative to all words. The proportion for each word found in drawer and non-drawer midterm responses was plotted on a log scale for the first (A) and second (B) midterm questions. Words closer to the diagonal line are words found in both drawer and non-drawer midterm responses. Words in gray are those with the highest word proportions, which represent some of the most common words found across all midterm responses. All of the midterm words between drawers and non-drawers are highly correlated.



131

Appendix S3.4: Sentiment values for AFINN, Bing et al., and NRC Emotion lexicons for drawer and non-drawer midterms. There was no statistical difference between the experimental and control groups for any of the three lexicons. For confidentiality, each student's name has been replaced by a participant number.



| **Lexicon** | AFINN | Bing et al. | NRC Emotion |
|---|---|---|---|
| **p-value** | 0.4884 | 0.6074 | 0.9843 |
| **mean (drawers)** | 15.6957 | -24.0000 | 3.9130 |
| **sd (drawers)** | 21.2505 | 18.2184 | 25.1720 |
| **mean (non-drawers)** | 11.6522 | -24.0870 | 7.6087 |
| **sd (non-drawers)** | 17.8367 | 10.5998 | 23.2178 |

Appendix S3.5: Quality of drawings were scored based on a range of seven levels of quality using criteria adapted from Nugraha 2018 and Reiss and Tunnicliffe 2001. There was no correlation between drawing quality and midterm scores within the experimental group (n = 23).

Level 1: No representation of any ecological concept

Level 2: One or more parts of an ecological concept placed at random

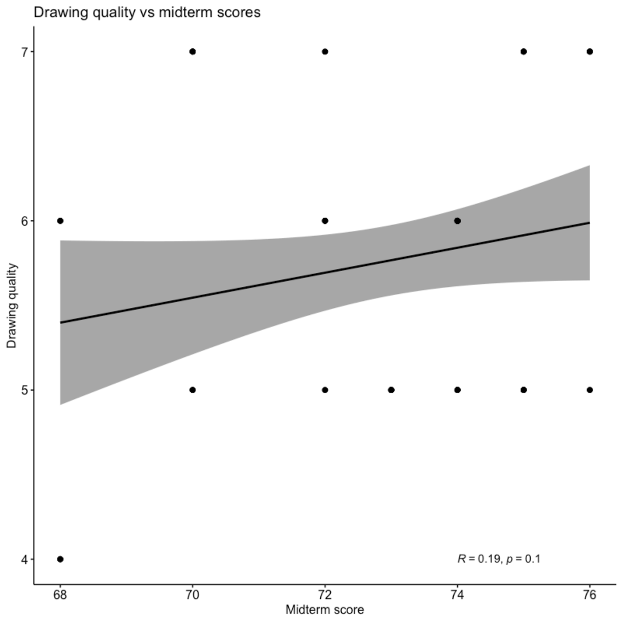
Level 3: One part of an ecological concept in appropriate position

Level 4: Two or more parts of an ecological concept in appropriate positions but no extensive relationships indicated between them

Level 5: One ecological concept indicated

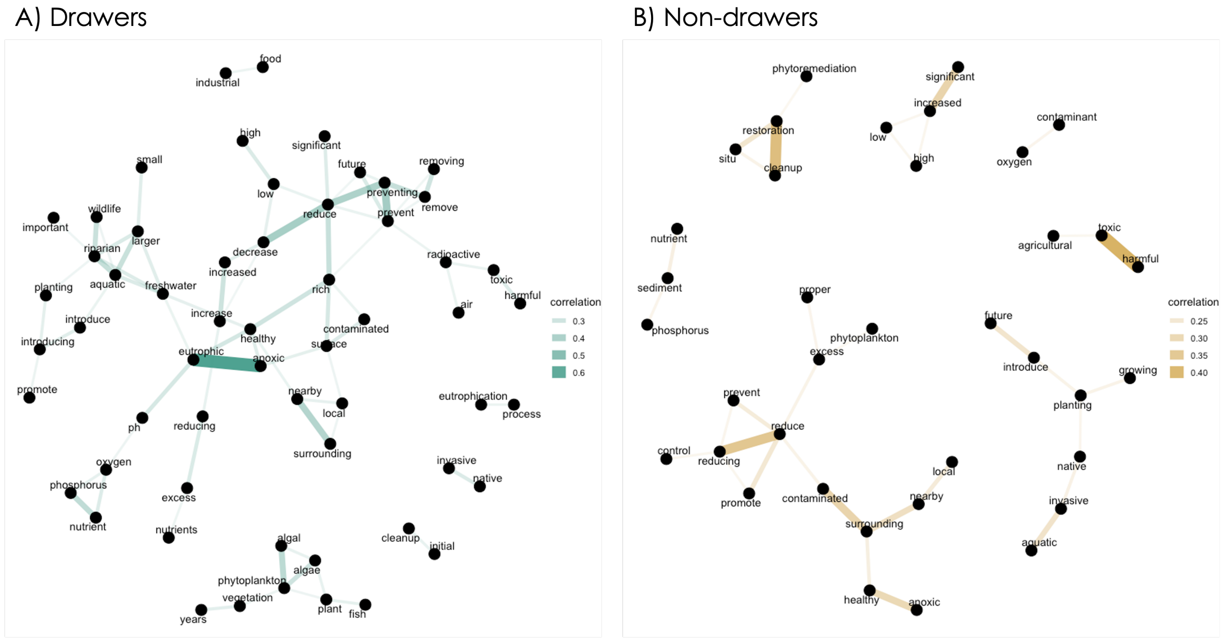
Level 6: Two or three ecological concepts indicated out of eutrophication, succession, assembly order, invasive species, ecosystem functions, health risks of contamination, or a restoration plan

Level 7: Comprehensive representation with four or more concepts indicated out of eutrophication, succession, assembly order, invasive species, ecosystem functions, health risks of contamination, or a restoration plan



References: 1) Nugraha, I. (2018). The use of drawing as an alternative assessment in biology teaching. Journal of Physics: Conference series, 1013-1018; 2) Reiss M J and Tunnicliffe S D 2001 Students' understandings of human organs and organ systems Research in Science Education 31 383–399.

Appendix S3.6: Correlation network analysis for paired words for responses in the entire midterm between drawers (A) and non-drawers (B). Drawers have a more complex and higher correlation between paired words compared to non-drawers. Note the difference in numeric scale within each group. This result aligns with that from the correlation network analysis for the first midterm question (see Figure 3.4).



## References

Ainsworth, S., & Loizou, A. T. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science*, *27*(4), 669–681.

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, *333*(6046), 1096–1097.

Ainsworth, S., Stieff, M., DeSutter, D., Tytler, R., Prain, V., Panagiotopoulos, D., Wigmore, P., Van Joolingen, W., Heijnes, D., & Leenaars, F. (2016). *Exploring the value of drawing in learning and assessment*. Singapore: International Society of the Learning Sciences.

Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education*, *53*(1), 74–85.

Armbruster, P., Patel, M., Johnson, E., & Weiss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE—Life Sciences Education*, *8*(3), 203–213.

Chang, C.-J., Liu, C.-C., & Tsai, C.-C. (2016). Supporting scientific explanations with drawings and narratives on tablet computers: An analysis of explanation patterns. *The Asia-Pacific Education Researcher*, *25*, 173–184.

Cheng, L., & Beal, C. R. (2020). Effects of student-generated drawing and imagination on science text reading in a computer-based learning environment. *Educational Technology Research and Development*, *68*(1), 225–247.

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105.

Coil, D., Wenderoth, M. P., Cunningham, M., & Dirks, C. (2010). Teaching the Process of Science: Faculty Perceptions and an Effective Methodology. *CBE—Life Sciences Education*, *9*(4), 524–535. https://doi.org/10.1187/cbe.10-01-0005

Connell, G. L., Donovan, D. A., & Chambers, T. G. (2016). Increasing the Use of Student-Centered Pedagogies from Moderate to High Improves Student Learning and Attitudes about Biology. *CBE—Life Sciences Education*, *15*(1), ar3. https://doi.org/10.1187/cbe.15-03-0062

Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction*, *9*(4), 343–363. https://doi.org/10.1016/S0959-4752(98)00051-6

De Bock, D., Verschaffel, L., Janssens, D., Van Dooren, W., & Claes, K. (2003). Do realistic contexts and graphical representations always have a beneficial impact on students' performance? Negative evidence from a study on modelling non-linear geometry problems. *Learning and Instruction*, *13*(4), 441–463.

De Smet, M. J. R., Broekkamp, H., Brand-Gruwel, S., & Kirschner, P. A. (2011). Effects of electronic outlining on students' argumentative writing performance. *Journal of Computer Assisted Learning*, *27*(6), 557–574. https://doi.org/10.1111/j.1365-2729.2011.00418.x

Dikmenli, M. (2010). Misconceptions of cell division held by student teachers in biology: A drawing analysis. *Scientific Research and Essay*, *5*(2), 235–247.

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, *22*(3), 293–331.

Ebert-May, D., Brewer, C., & Allred, S. (1997). Innovation in large lectures: Teaching for active learning. *Bioscience*, *47*(9), 601–607.

Federer, M. R., Nehm, R. H., & Pearl, D. K. (2016). Examining Gender Differences in Written Assessment Tasks in Biology: A Case Study of Evolutionary Explanations. *CBE—Life Sciences Education*, *15*(1), ar2. https://doi.org/10.1187/cbe.14-01-0018

Fiorella, L., & Mayer, R. E. (2016). Effects of observing the instructor draw diagrams on learning from multimedia messages. *Journal of Educational Psychology*, *108*(4), 528–546. https://doi.org/10.1037/edu0000065

Fiorella, L., & Zhang, Q. (2018). Drawing boundary conditions for learning by drawing. *Educational Psychology Review*, *30*, 1115–1137.

Freeman, S. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Freeman, S., O'Connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., Dirks, C., & Wenderoth, M. P. (2007). Prescribed active learning increases performance in introductory biology. *CBE—Life Sciences Education*, *6*(2), 132–139.

Gobert, J. D., & Clement, J. J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *36*(1), 39–53.

Goubeaud, K. (2010). How is Science Learning Assessed at the Postsecondary Level? Assessment and Grading Practices in College Biology, Chemistry and Physics. *Journal of*

*Science Education and Technology*, *19*(3), 237–245. https://doi.org/10.1007/s10956-009-9196-9

Guilford, W. H. (2001). Teaching Peer Review and the Process of Scientifc Writing. *Advances in Physiology Education*, *25*(3), 167–175. https://doi.org/10.1152/advances.2001.25.3.167

Hackling, M. W., & Prain, V. (2005). *Primary connections: Stage 2 trial*.

Hailikari, T., Katajavuori, N., & Lindblom-Ylanne, S. (2008). The relevance of prior knowledge in learning and instructional design. *American Journal of Pharmaceutical Education*, *72*(5).

Heideman, P. D., Flores, K. A., Sevier, L. M., & Trouton, K. E. (2017). Effectiveness and adoption of a drawing-to-learn study tool for recall and problem solving: Minute sketches with folded lists. *CBE—Life Sciences Education*, *16*(2), ar28.

Hensley, L., Kulesza, A., Peri, J., Brady, A. C., Wolters, C. A., Sovic, D., & Breitenberger, C. (2021). Supporting Undergraduate Biology Students' Academic Success: Comparing Two Workshop Interventions. *CBE—Life Sciences Education*, *20*(4), ar60. https://doi.org/10.1187/cbe.21-03-0068

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177. https://doi.org/10.1145/1014052.1014073

Jenkinson, J. (2018). Molecular biology meets the learning sciences: Visualizations in education and outreach. *Journal of Molecular Biology*, *430*(21), 4013–4027.

Jerde, C. L., & Taper, M. L. (2004). *Preparing Undergraduates for Professional Writing: Evidence Supporting the Benefits of Scientific Writing within the Biology Curriculum*. *33*(7), 34–37.

Justice, C., Rice, J., Warry, W., Inglis, S., Miller, S., & Sammon, S. (2006). Inquiry in Higher Education: Reflections and Directions on Course Design and Teaching Methods. *Innovative Higher Education*, *31*(4), 201–214. https://doi.org/10.1007/s10755-006-9021-9

Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 355–365. https://doi.org/10.1037/0278-7393.14.2.355

Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, *1*(1), 1–26. https://doi.org/10.17239/jowr-2008.01.01.1

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, *12*(1), 1–10. https://doi.org/10.1016/S0959-4752(01)00014-7

Knight, J. K., & Wood, W. B. (2005). Teaching more by lecturing less. *Cell Biology Education*, *4*(4), 298–310.

Korakakis, G., Pavlatou, E. A., Palyvos, J. A., & Spyrellis, N. (2009). 3D visualization types in multimedia applications for science learning: A case study for 8th grade students in Greece. *Computers & Education*, *52*(2), 390–401.

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, *13*(2), 205–226. https://doi.org/10.1016/S0959-4752(02)00021-X

Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior*, *25*(2), 284–289. https://doi.org/10.1016/j.chb.2008.12.010

Liang, J.-C., & Tsai, C.-C. (2010). Learning through science writing via online peer assessment in a college biology course. *The Internet and Higher Education*, *13*(4), 242–247. https://doi.org/10.1016/j.iheduc.2010.04.004

Lin, L., Lee, C. H., Kalyuga, S., Wang, Y., Guan, S., & Wu, H. (2017). The effect of learner-generated drawing and imagination in comprehending a science text. *The Journal of Experimental Education*, *85*(1), 142–154.

Martin, A. J., Wilson, R., Liem, G. A. D., & Ginns, P. (2013). Academic momentum at university/college: Exploring the roles of prior learning, life experience, and ongoing performance in academic achievement across time. *The Journal of Higher Education*, *84*(5), 640–674.

McClean, P., Johnson, C., Rogers, R., Daniels, L., Reber, J., Slator, B. M., Terpstra, J., & White, A. (2005). Molecular and cellular biology animations: Development and impact on student learning. *Cell Biology Education*, *4*(2), 169–179.

McElhaney, K. W., Chang, H.-Y., Chiu, J. L., & Linn, M. C. (2015). Evidence for effective uses of dynamic visualisations in science curriculum materials. *Studies in Science Education*, *51*(1), 49–85.

Mohammad, S., & Turney, P. (2010). *Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon*.

Mohammad, S., & Turney, P. D. (2013). *NRC Emotion Lexicon*.

Mostrom, A. M., & Blumberg, P. (2012). Does Learning-Centered Teaching Promote Grade Improvement? *Innovative Higher Education*, *37*(5), 397–405. https://doi.org/10.1007/s10755-012-9216-1

Newton, X. A., & Tonelli, E. P. (2020). Building undergraduate STEM majors' capacity for delivering inquiry-based mathematics and science lessons: An exploratory evaluation study. *Studies in Educational Evaluation*, *64*, 100833. https://doi.org/10.1016/j.stueduc.2019.100833

Nielsen, F. Å. (2011). *A new ANEW: Evaluation of a word list for sentiment analysis in microblogs* (arXiv:1103.2903). arXiv. http://arxiv.org/abs/1103.2903

Nugraha, I. (2018). The use of drawing as an alternative assessment tool in biology teaching. *Journal of Physics: Conference Series*, *1013*(1), 012016.

Peart, D. J. (2022). Hand drawing as a tool to facilitate understanding in undergraduate human biology: A critical review of the literature and future perspectives. *Studies in Science Education*, *58*(1), 81–93.

Prain, V., & Tytler, R. (2013). Learning through the affordances of representation construction. In *Constructing representations to learn in science* (pp. 67–82). Brill.

Quillin, K., & Thomas, S. (2015). Drawing-to-learn: A framework for using drawings to promote model-based reasoning in biology. *CBE—Life Sciences Education*, *14*(1), es2.

Ryoo, K., & Linn, M. C. (2012). Can dynamic visualizations improve middle school students' understanding of energy in photosynthesis? *Journal of Research in Science Teaching*, *49*(2), 218–243.

Sandrone, S., Scott, G., Anderson, W. J., & Musunuru, K. (2021). Active learning-based STEM education for in-person and online learning. *Cell*, *184*(6), 1409–1414.

Schmeck, A., Mayer, R. E., Opfermann, M., Pfeiffer, V., & Leutner, D. (2014). Drawing pictures during learning from scientific text: Testing the generative drawing effect and the prognostic drawing effect. *Contemporary Educational Psychology*, *39*(4), 275–286.

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *46*(6), 632–654.

Sebesta, A. J., & Bray Speth, E. (2017). How should I study for the exam? Self-regulated learning strategies and achievement in introductory biology. *CBE—Life Sciences Education*, *16*(2), ar30.

Silge, J., & Robinson, D. (2016). tidytext: Text Mining and Analysis Using Tidy Data Principles in R. *The Journal of Open Source Software*, *1*(3), 37. https://doi.org/10.21105/joss.00037

Stieff, M. (2017). Drawing for promoting learning and engagement with dynamic visualizations. *Learning from Dynamic Visualization: Innovations in Research and Application*, 333–356.

Thompson, R. A., & Zamboanga, B. L. (2004). Academic aptitude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology*, *96*(4), 778.

Torrance, M., Thomas, G. V., & Robinson, E. J. (2000). Individual differences in undergraduate essay-writing strategies: A longitudinal study. *Higher Education*, *39*(2), 181–200. https://doi.org/10.1023/A:1003990432398

Udovic, D., Morris, D., Dickman, A., Postlethwait, J., & Wetherwax, P. (2002). Workshop biology: Demonstrating the effectiveness of active learning in an introductory biology course. *Bioscience*, *52*(3), 272–281.

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, *31*(2), 142–166.

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review*, *17*, 285–325.

Walker, J. D., Cotner, S. H., Baepler, P. M., & Decker, M. D. (2008). A delicate balance: Integrating active learning into a large lecture course. *CBE—Life Sciences Education*, *7*(4), 361–367.

Wright, C. D., Eddy, S. L., Wenderoth, M. P., Abshire, E., Blankenbiller, M., & Brownell, S. E. (2016). Cognitive Difficulty and Format of Exams Predicts Gender and Socioeconomic Gaps in Exam Performance of Students in Introductory Biology Courses. *CBE—Life Sciences Education*, *15*(2), ar23. https://doi.org/10.1187/cbe.15-12-0246

Wu, S. P. W., & Rau, M. A. (2019). How Students Learn Content in Science, Technology, Engineering, and Mathematics (STEM) Through Drawing Activities. *Educational Psychology Review*, *31*(1), 87–120. https://doi.org/10.1007/s10648-019-09467-3

Yarden, H., & Yarden, A. (2010). Learning using dynamic and static visualizations: Students' comprehension, prior knowledge and conceptual status of a biotechnological method. *Research in Science Education*, *40*, 375–402.