# DIALECTS OVER SPACE AND TIME:
# CULTURAL IDENTITY AND EVOLUTION IN SPERM WHALE CODAS

by

Taylor A. Hersh

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
November 2021

*To my parents, who have always been the wind behind my sails,*

*And to Jonah, for adding a little magic to the journey.*

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ABSTRACT

The overarching goal of my thesis is to better understand how sperm whale (*Physeter macrocephalus*) dialects vary over space and time, and to quantify the implications of these dynamics on cultural identity and evolution in sperm whale vocalizations. Female and immature sperm whales live in matrilineal social units that are stable over time, and different social units only associate if they belong to the same cultural clan. Clans forage, move, dive, associate, and distribute differently, but the primary way we distinguish them is through their unique dialects. These dialects are comprised of stereotyped patterns of clicks, called codas, and my thesis seeks to clarify the interplay between these social vocalizations and cultural identity in sperm whale clans. Inspired by rhythm in codas, I first survey the literature for evidence of rhythm in cetacean vocalizations and use a comparative lens to theorize about the functional roles it plays. I show that rhythm is common in cetacean vocalizations, but that it may be used in different behavioral contexts by mysticetes and odontocetes. In sperm whales, vocal rhythm is apparent across contexts, suggesting that it is a fundamental feature of communication. Next, I introduce a new method (IDcall) that detects putative biological structure in acoustic datasets using characteristic, repeated call types. I show that IDcall's underlying theory is broadly applicable and can be used to detect sperm whale clans from codas, wren subspecies from songs, and cricket species from songs. Using acoustic data from 25 locations spanning 42 years, I then investigate the spatial and temporal dynamics of sperm whale clan dialects. In the spatial domain, I document the presence and distribution of seven clans in the Pacific Ocean, and provide empirical evidence that certain coda types function as symbolic markers of sperm whale clan identity, like ethnic markers in humans. In the temporal domain, I leverage long-term research efforts in the Mediterranean, eastern Caribbean, and Galápagos Islands to show that the fine-scale structure of coda types can change over decadal timespans within clans. Collectively, these findings illustrate the interplay between vocalizations and sociality in sperm whales, and emphasize the need for a global, multi-cultural approach to study and conserve this global, multi-cultural species.

# LIST OF ABBREVIATIONS AND SYMBOLS USED

| | |
|---|---|
| $\mathrm{acsim}_{\mathrm{allbt}}$ | between-clan acoustic similarity calculated using all codas |
| $\mathrm{acsim}_{\mathrm{allwi}}$ | within-clan acoustic similarity calculated using all codas |
| $\mathrm{acsim}_{\mathrm{diffbt}}$ | between-clan acoustic similarity calculated as ($\mathrm{acsim}_{\mathrm{IDbt}} - \mathrm{acsim}_{\mathrm{nonIDbt}}$) |
| $\mathrm{acsim}_{\mathrm{diffwi}}$ | within-clan acoustic similarity calculated as ($\mathrm{acsim}_{\mathrm{IDwi}} - \mathrm{acsim}_{\mathrm{nonIDwi}}$) |
| $\mathrm{acsim}_{\mathrm{IDbt}}$ | between-clan acoustic similarity calculated using identity codas |
| $\mathrm{acsim}_{\mathrm{IDwi}}$ | within-clan acoustic similarity calculated using identity codas |
| $\mathrm{acsim}_{\mathrm{nonIDbt}}$ | between-clan acoustic similarity calculated using non-identity codas |
| $\mathrm{acsim}_{\mathrm{nonIDwi}}$ | within-clan acoustic similarity calculated using non-identity codas |
| AIC | Akaike Information Criterion |
| AICc | Akaike Information Criterion corrected for small sample size |
| AtPAN | Atlantic coast of Panama |
| BAK | Baker Island |
| BAL | Balearic Islands |
| BC-hil | backcross between Hil and F1 |
| BC-leu | backcross between Leu and F1 |
| BIC | Bayesian Information Criterion |
| BOW | Bowie Seamount |
| °C | degrees Celsius |
| *c* | an identity clade (in IDcall) |
| CAR | Caribbean |
| CHL | Chile |
| CHL_N | Northern Chile |
| CHL_S | Southern Chile |
| $\mathrm{clan}_{\mathrm{allyears}}$ | clan-level analysis using codas from all years |
| $\mathrm{clan}_{\mathrm{largeyears}}$ | clan-level analysis using codas from years with ≥25 codas |
| Com | *Teleogryllus commodus* |
| *critfact* | factor by which a given call type's usage must be higher in one clade than in any other clade for it to be considered an identity call (in IDcall) |
| CV | coefficient of variation |
| D | decreasing intervals between coda clicks |

| | |
|---|---|
| DNA | deoxyribonucleic acid |
| EAS | Easter Island |
| ECU | Ecuador |
| EC1 | Eastern Caribbean 1 clan |
| EC2 | Eastern Caribbean 2 clan |
| ESP | Equatorial South Pacific |
| ESU | Evolutionarily significant unit |
| F1 | first generation hybrid |
| FP | Four-Plus clan |
| GAL | Galápagos Islands |
| GOM | Gulf of Mexico |
| Hil | parental genotype *Henicorhina leucophrys hilaris* |
| Hz | hertz |
| $i$ | a call (in IDcall) |
| I | increasing intervals between coda clicks |
| ICI | inter-click interval |
| ICL | Integrated Completed Likelihood |
| $icp$ | clade identity call proportion (in IDcall) |
| ID | identity |
| IDcall | the identity call method |
| IUCN | International Union for Conservation of Nature |
| $j$ | a call type (in IDcall) |
| JAR | Jarvis Island |
| JPN_K | Kumano coast of Japan |
| JPN_O | Ogasawara Islands of Japan |
| kHz | kilohertz |
| KIR | Kiribati |
| km | kilometer |
| LAN | Lesser Antilles |
| Leu | parental genotype *Henicorhina leucophrys* |
| m | meter |

| | |
|---|---|
| MED | Mediterranean clan |
| MID | Midway Atoll |
| min | minutes |
| *minrep* | minimum number of repertoires required for an identity clade to form (in IDcall) |
| MNP | Mariana Islands |
| MRQ | Marquesas Islands |
| ms | milliseconds |
| n | sample size |
| *N* | number of quantitative measures (in IDcall) |
| NA | not applicable |
| NaN | not a number |
| NRU | Nauru |
| NZL | New Zealand |
| NZL_N | Northern New Zealand |
| NZL_S | Southern New Zealand |
| Oce | *Teleogryllus oceanicus* |
| Out | outlier |
| p | p-value |
| PacPAN | Pacific coast of Panama |
| PAL | Palau |
| PALI | Palindrome clan |
| PAN | Panama |
| PC | principal component |
| PCA | principal component analysis |
| PER | Peru |
| PNG | Papua New Guinea |
| PO | Plus-One clan |
| *r* | a repertoire (in IDcall) |
| R | regular (i.e. isochronous) spacing between coda clicks |
| REG | Regular clan |

| | |
|---|---|
| RI | Rapid Increasing clan |
| s | seconds |
| S | supplemental |
| SD | standard deviation |
| SE | standard error |
| SH | Short clan |
| SHE | Short East clan |
| SHW | Short West Clan |
| SI | Slow Increasing clan |
| SOC | Sea of Cortez |
| TON | Tonga |
| $U$ | call type usage, ranging from 0 to 1 (in IDcall) |
| VEE | variable volume, equal shape, equal orientation |
| VEV | variable volume, equal shape, variable orientation |
| vs | versus |
| VVE | variable volume, variable shape, equal orientation |
| VVV | variable volume, variable shape, variable orientation |
| + | extended interval between coda clicks |

# ACKNOWLEDGEMENTS

“Love is a house with many rooms…”
—Yann Martel

The past five years have been the happiest of my life, due in no small part to the people I was surrounded by. To my supervisor and mentor, Hal Whitehead, thank you for what I can only describe as a dream of a Ph.D. Your patience, humor, grit, and drive are constant sources of inspiration to me. Thank you for always having time for a call; for sending me into the field early and often; for embracing my quirks; for pushing me when I needed it, but also granting me the independence to grow as a researcher. Being your student has meant non-stop opportunities to learn from you, both in the lab and in the field, and I look forward to many more years of working together. Getting to know and spend time with Jennifer and the pups was a lovely bonus of being your student.

To Shane Gero, mentor extraordinaire and birthday twin: it is impossible to quantify the impact you have had on me, both as a scientist and as a person. You were helping me before I even became a graduate student (editing the very funding application that allowed me to come to Dalhousie), and you haven’t stopped since. Based on my records, we have spent at least 86.2 hours on video calls over the years (yes, I kept track!), which is to say nothing of the hundreds of hours we spent together in the field. Thank you for always believing in me, even when (especially when) I didn’t believe in myself. I feel so lucky to be part of the DSWP and will forever be brainstorming ways to poco poco my way back to the island and the whales with you.

The whales were an endless source of inspiration and comfort throughout my degree. I feel so privileged to have had a small window into their lives, and to have spent so many hours eavesdropping on their conversations. Over the years, many people have asked me when I will stop marking codas. I’ve come to realize that the answer is never, and I wouldn’t have it any other way.

What can I possibly say to my family except thank you? I am incredibly lucky to have supportive parents who have gifted me drive, courage, and independence in equal measure. Your faith in me has slowly but surely bolstered my faith in myself, and this thesis is as much a reflection of you as it is of me. All my love, always.

To my committee members, Andy Horn and Aaron Newman, thank you for being enthusiastic about this work, always forwarding interesting papers, and asking the good, hard questions when I needed them. To my external examiner, Stephanie King, thank you for reading and providing feedback on my work; I have admired your research for years and it was an honor to have you read mine.

To Laura Joan Feyrer: there is a beautiful symmetry in the fact that I started and ended my Ph.D. by helping you prep for a field season. Thank you for still being my friend even though I didn't tell you about my racoon hallucinations immediately. There is no one I would rather call than you to help me respool messy reel-to-reel tapes (or my life).

I was lucky enough to be adopted by not one but two fantastic families during my time in Halifax. To the Ali family—Jasmine, Marty, Sylvia, and Sam—thank you for welcoming me with open arms into your home/cabin/hair salon/convenience store/print shop/lives. Jasmine, my soulmate of a friend, my life is so much brighter with you in it. To the Youssef/Aly/McCain family—Amani (Mamani), Ameto Sanaa, Alaa, Khlood, Khaled, Lamis, Maged, Naima, Noor, Omar, Sara, and Scott—thank you for filling my nights with laughter, Catan, dancing, and food. Noor, living together (with Sasha) for the last year of our degrees was a dream come true. You are present in every page of this thesis in big and small ways, from your always helpful edits and research advice to your never-ending supply of hugs and pep talks. Scott (Scoot, Scooter, Scotty), thank you for reminding me to not take life too seriously and for always helping me reason through sticky/anxiety-inducing situations. Noor, Scott, and Omar are the reason I survived lockdown #2; Travis Aten, Mark Devereux, Franky, and David are the reason I survived lockdown #1. I love you all so much.

I made many fantastic friends through the Biology Organization of Graduate Students, including Benia Nowak, Megan Roberts, Cait Nemeczek, Cat Bannon, Joceline Lebans, Hilary Mann, Georgia Denbigh, Nathan Hoffman, Lisette Delgado, Loay Jabre, Sara Ericsson, Mili Sanchez, and more. You are the best and brightest. My lab mates—Wilfried Beslin, Ana Eguiguren, Laura Joan Feyrer, Christine Konrad, Katie Kowarski, Julia Riley, Félicia Vachon, and Elizabeth Zwamborn—were a steady source of camaraderie and support over the years.

My Dominica field family/wolf pack will always hold a special place in my heart. To the *Flying Fish* crew—Dave Fabien, Jaey Shillingford, Ricardo Barrett, Kevin George, Candy Stoute, and Zethra Baron—thank you for teaching me the ropes and always making me laugh. I love you all plenty a lot. I am a better scientist and person for having worked with Pernille Tønnesen, who is the most level-headed, light-hearted researcher I know (even if she does use too much electrical tape). I would not have wanted to island hop from Dominica to Trinidad with anyone but Ellen Jacobs; thank you for being a squirrel.

To Luke Rendell: I will always be grateful for the Caribbean survey we did together, and for your offhand comment about how I should come visit you in Scotland. Thank you for hosting me at the University of St. Andrews for what became one of my favorite semesters of graduate school. You gave me a change of pace and place when I needed it most. Without my trip to Scotland, I never would have met Jack Rayner, Sam Sturiale, and Tom Hitchcock, who are now three of my very best friends. I love you to Wewak/Kokomo/Bungalow and back, you weirdos—thanks for keeping me (in)sane.

I would be lost in life without my two oldest friends, Olga Graves and Paige Byrns. You two know me better than anyone and have been my rocks throughout this degree. We have been cheering each other on since childhood, and I feel ready for the future, whatever it holds, because I have you both by my side. Us three, until the sun dies.

To the members of the Comparative Bioacoustics Group at the Max Planck Institute for Psycholinguistics—Andrea Ravignani, Koen de Reus, Yannick Jadoul, Laura Verga, Silvia Leonetti, Diandra Düngen, Masha Goncharova, Mary Anichini, Jelle van der Werff, and Christoforos Souganidis—thank you for making an ending feel more like a beginning. Your friendship, advice, good humor, and helpful edits were much appreciated in the final leg of my degree.

I am grateful to the various organizations that funded my research and fieldwork throughout my degree, including the American Philosophical Society, the Animal Behavior Society, the Arizona Center for Nature Conservation, the Brevard Zoo Conservation Fund, Dalhousie University, the Explorer's Club, the Killam Trust, Mitacs Globalink, National Geographic, the Natural Sciences and Engineering Research Council of Canada, the Nova Scotian government, PADI Foundation, and the Women Divers Hall of Fame.

Many undergraduate students volunteered their time to help me audit acoustic files, mark sperm whale codas, and transcribe photo-identification datasheets. I am especially indebted to Courtney Baumgartner, Kathleen Buffet, Michaela Ermanovics, Stephen Feng, Yaly Mevorach, and Eliza-Jane Morin for their help.

It takes a clan to write a thesis. I will always be thankful for the people and whales in mine.

# CHAPTER 1

# INTRODUCTION

> "The nature of the voyage is set before you cast off. A sea passage is shaped by the boat's time attached to the land. Every moment at sea is dependent on, and even twinned to, a moment in harbour. What a boat sails on and in is not only the ocean and the wind, but the days, weeks, and months tied up alongside."
>
> —Adam Nicolson

Communication—the sharing of information between self and other—can be observed in every major branch of the tree of life (including trees themselves; Simard, 2018). For some taxa, it is the foundation on which complexity (be it social, behavioral, etc.) relies. The most familiar examples include our own communication system and those of other animals we consciously encounter in our day-to-day lives, including various insects (e.g. cicadas; Fonseca, 2014); birds (e.g. zebra finches, *Taeniopygia guttata*; Elie & Theunissen 2016), and mammals (e.g. squirrels; Diggins, 2021). However, communication also occurs at microscopic scales. A marine bacterium, *Vibrio harveyi*, produces cell-to-cell signaling molecules that trigger bioluminescence (Waters & Bassler, 2005). This process is known as quorum sensing and muddles "the distinction between prokaryotes and eukaryotes because it enables bacteria to act as multicellular organisms" (Waters & Bassler, 2005). The communication systems of various slime molds have improved swarm algorithms (Varughese et al., 2019), inspired wireless sensor network routing systems (Li et al., 2011), and captivated public attention (Fox, 2021; Jabr, 2012). Communication thus seems to be a hallmark of life, but the preferred means of communicating can vary across species and environments.

## 1.1 – Modalities of communication

Communication can occur via multiple sensory modalities. For example, honey bees use a visual display, known as the waggle dance, to share information on the location of food with nest-mates (Frisch, 2013; I'Anson Price & Grüter, 2015). In the chemical domain, various anurans (i.e. frogs and toads) find potential mates by tracking the release of pheromones through water (reviewed in Belanger & Corkum, 2009). Several species of freshwater and marine fish use electric signals for social communication (Hopkins, 1988), and tactile behaviors, like grooming and embracing, are key to maintaining social structure and relieving stress in chimpanzee (*Pan troglodytes*) communities (reviewed in Hertenstein et al., 2006). As a bioacoustician, I have always been most fascinated by acoustic communication, in which acoustic signals are used to transmit information. Acoustic communication is used by a diversity of species in a diversity of contexts, including foraging (e.g. meerkats, *Suricata suricatta*; Gall & Manser, 2017), courtship (e.g. satin bowerbirds, *Ptilonorhynchus violaceus*; Loffredo & Borgia, 1986), warning (e.g. vervet monkeys, *Cercopithecus aethiops*; Seyfarth et al., 1980), and socializing (e.g. Pacific humpback dolphins, *Sousa chinensis*; Van Parijs & Corkeron, 2001).

When acoustic communication is carried out via vocalizations[1], it can be called vocal communication. Despite the rapidity with which we can accumulate acoustic data and the sheer abundance of studies on acoustic signaling, there are still challenges inherent in understanding vocal communication systems. As Sainburg et al. (2020) aptly put it, "The characterization and abstraction of vocal communication signals remains both an art and a science."

## 1.2 – Describing vocal communication: an art and a science

Much of the early vocal communication research was done in humans (*Homo sapiens*) and songbirds and focused on determining whether vocalizations were innate or socially learned (an offshoot of the polarizing nature vs. nurture debate) (Doupe & Kuhl,

---

[1] Throughout this thesis, I use the word 'vocalization' to mean any acoustic call produced by an animal, regardless of whether the call is produced via the vocal tract/cords (as some definitions of 'vocalization' require) or via other anatomical features (e.g. the phonic lips in odontocetes). 'Call' and 'vocalization' are used interchangeably.

1999; Keller, 2010; Nieder & Mooney, 2020). It has become increasingly clear, however, that many vocalizations have both innate and social components (Arriaga & Jarvis, 2013; Egnor & Hauser, 2004; Knörnschild, 2014; Lipkind et al., 2020). With the right combination of neural underwiring and sensory input from tutors, most animals can acquire the typical vocal repertoires of adult conspecifics (Doupe & Kuhl, 1999; Lipkind et al., 2020). In a similar vein, species have historically been classified as either vocal or non-vocal learners, but research suggests that vocal learning abilities fall along a gradient (Arriaga et al., 2012; Petkov & Jarvis, 2012). At a more fundamental level, vocalizations themselves have often been classified as either discrete or graded, but this too may be better conceptualized as a continuum (Cusano et al., 2021; Lipkind et al., 2020; Sainburg et al., 2020). Even the distinctions between "notes" and "calls", or between "good" and "bad" quality vocalizations based on signal-to-noise ratios, can be difficult to quantify (Fournet et al., 2018; Weirathmueller et al., 2017). Despite the human affinity for binarism, these examples emphasize that spectra are often more appropriate than dichotomies when describing vocal communication systems.

Accepting that boundaries in vocal communication systems and abilities can be fuzzy does not negate the fact that vocalizations for most taxa are not randomly distributed in multivariate space (the dimensions of which are typically limited by vocal anatomy, which imposes sound production constraints; e.g. Lieberman et al., 1969; Reby & McComb, 2003; De Boer, 2010). For example, five-click codas (discussed in detail in section 1.5) made by eastern Caribbean sperm whales form distinct clusters, not an amorphous blob, in multivariate space (Figure 1.1). How and why do distinct vocalizations emerge within a population?

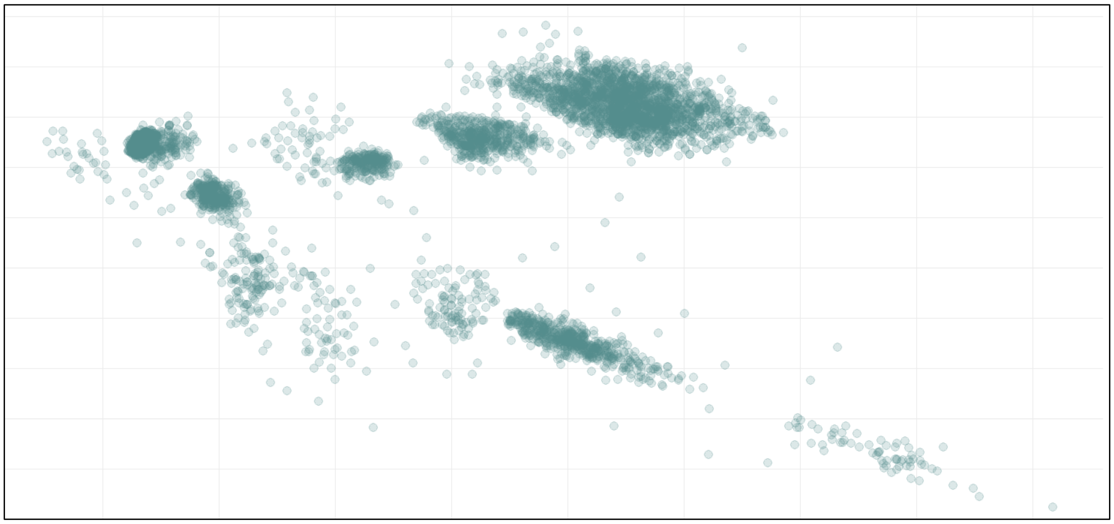


*Figure 1.1* – Multivariate representation of the inter-click intervals for five-click codas made by eastern Caribbean sperm whales. Each point represents a single coda (n=5,134). See section 1.5 for more details on codas.

Different vocalizations can be produced by varying spectral (e.g. fundamental frequency, bandwidth) and/or temporal (e.g. rhythm, tempo) parameters, and these vocalizations can subsequently play different roles and convey different types of information in a communication system. Expressivity—"the capacity of a signaling system to convey different meanings" (Ravignani & Madison, 2017)—is often directly related to the diversity of vocalization types (e.g. humans, Bentz et al., 2017; spectacled parrotlets, *Forpus conspicillatus,* Wanker et al., 2005; pied babblers, *Turdoides bicolor*, Engesser et al., 2016). Cusano et al. (2019) hypothesized that calls that fall more towards the discrete (i.e. stereotyped) end of the discrete/graded continuum typically provide fixed information, such as sex, body size, age, or identity—signaler features that are generally stable over (at least) short timescales. Such calls are generally consistent within and between behavioral contexts and form denser, more discrete cores in multivariate space (Cusano et al., 2019). It is worth noting here that the 'identity' information contained in certain vocalizations can be individual (e.g. bottlenose dolphins, *Tursiops truncatus;* Janik et al., 2006), species (e.g. common dolphins, *Delphinus* spp.; Oswald et al., 2021), population (e.g. swamp sparrows, *Melospiza georgiana;* Balaban, 1988), and/or cultural (e.g. naked mole-rats, *Heterocephalus glaber;* Barker et al., 2021). In contrast, calls that fall more towards the graded (i.e. variable) end of the continuum often provide information on internal motivational state or arousal level—traits that can change rapidly over short timescales (Cusano et al., 2019). For example, graded calls can convey information about perceived threats (Coss et al., 2007; Schehka et al., 2007) or aggression levels (DuBois et al., 2009).

Social systems are formed by information moving between individuals, and distinct vocalizations are one way for individuals to convey different types of information to others. Communication can thus be thought of as the glue of sociality—defined here as the degree to which individuals in a population live together, interact, and cooperate (Wey et al., 2008). Conversely, without sociality, the need for a shared communication system is largely negated.

### 1.3 – Sociality can augment vocalizations, vocalizations can augment sociality

The social complexity hypothesis for communication posits that groups of animals with greater social complexity will also demonstrate greater complexity in their communication systems compared to groups with less complex social structures (Freeberg et al., 2012). The rationale is that individuals who live in complex societies typically have to manage many individual interactions and/or types of relationships, which can be facilitated through a more complex communication system (Freeberg et al., 2012). Evidence supporting this hypothesis has been documented in many taxa, including bats (Knörnschild et al., 2020; Wilkinson, 2003), non-human primates (Bouchet et al., 2013; McComb & Semple, 2005), birds (Freeberg, 2006; Krams et al., 2012), squirrels (Blumstein & Armitage, 1997), and marmots (Blumstein & Armitage, 1997).

However, while the social complexity hypothesis for communication is explicitly framed with social complexity *causing* vocal complexity, that directionality has been questioned (McComb & Semple, 2005; Ord & Garcia-Porta, 2012). Having a complex communication system enables many finely graded messages to be sent and received, which can add more dimensions to how animals interact and associate. McComb and Semple (2005) point out that while both group size and time spent grooming are strongly positively associated with vocal repertoire size in primates, other selective forces may have driven vocal complexity evolution, which then facilitated larger and more complex social structures. A feedback loop likely exists between the two, with sociality augmenting vocalizations and vocalizations augmenting sociality in significant ways. In dynamic ocean environments, acoustic communication through vocalizations is extremely common and effective, making the ocean a fascinating habitat to further investigate social/vocal interplays.

### 1.4 – Communication in the sonic sea

Features of the ocean limit the range of many communication modalities that are used effectively on land. Sunlight can only strongly penetrate the upper ~200 m of the water column (with no light reaching 1 km or deeper), which means that most visual displays are largely restricted to shallower waters (but see Haddock et al., 2010 for a review of ocean bioluminescence). Water turbidity further limits the range at which such displays can be detected, making them useful for short-range but not long-range communication. Chemical communication is typically short-range as well, given that chemical cues are released into a chemically complex background and can diffuse rapidly in water (although odor plumes from tuna schools and whale carcasses can encompass hundreds or thousands of meters) (Atema, 1995). The strength of electric signals produced in water decreases rapidly with increasing distance from the source, limiting the reach of such signals (Hopkins, 2009), and tactile communication, by definition, requires extreme proximity of individuals. Indeed, the modality best suited to long-range communication in the ocean is acoustic signaling because water retains sound energy quite well (especially compared to air). Exactly how far a sound wave propagates, however, is affected by three processes: refraction, reflection, and absorption.

Refraction occurs when sound energy is diverted from its normal path by changes in its speed of propagation. This can be affected by temperature, salinity, and pressure, with sound speeds increasing in hotter, saltier, and deeper waters. When sound energy is redirected due to a change in media, it is called reflection. In the ocean, a sound wave can reflect off any mass in the water column, as well as off the seafloor and the surface, resulting in some of the sound energy being scattered. Finally, absorption happens when sound energy is converted into thermal energy. High frequency sounds are absorbed more quickly than low frequency sounds. Animals that communicate using low frequency sounds, such as blue whales (Širović et al., 2007), can thus communicate over much larger distances than animals that use high frequency sounds, such as porpoises (Hansen et al., 2008). While the amount of absorption largely depends on a sound's frequency, it can also be affected by temperature, salinity, and pressure, such that different oceans absorb sound differently (Ainslie & McColm, 1998).

Unsurprisingly, many marine animals use sound to communicate, including invertebrates, fish, and cetaceans (reviewed in Tyack, 1998). The bulk of my thesis work has focused on how one cetacean in particular, the sperm whale (*Physeter macrocephalus*), uses sound.

### 1.5 – Sperm whales as a case study: lessons from an ocean leviathan

A quick Google search for the term "whale icon" results in a barrage of images that closely resemble sperm whales. Abstracted away from the real thing, most people do not realize that they are looking at a caricature of the legendary leviathan, with a boxy head, small flippers, and a thin jaw. Through decades of research, we have been able to transform that rough caricature into an intimate portrait.

Sperm whales have occupied a unique place in the zeitgeist since the whaling heyday of the 19th century (Bannister et al., 2012). This position was further cemented by the 1851 publication (and belated appreciation) of Herman Melville's masterpiece, *Moby Dick*, as well as the discovery that sperm whales are the only known predator of the almost mythical giant (*Architeuthis dux*) and colossal (*Mesonychoteuthis hamiltoni*) squids (Clarke, 1980; Remeslo et al., 2019; Roper & Boss, 1982). What became clear over a century after *Moby Dick* was published is that while Melville's characterization of sperm whales as vengeful, monstrous killing machines had whiffs of accuracy for animals trying to escape the sharp harpoons of whalers (Chase, 2015; Philbrick, 2001), it neglected the largely peaceful nature of these gentle giants (Beale, 1839).

The sperm whale belongs to a phylogenetically distinct cetacean clade, branching away from the rest of the tree (including other odontocetes, i.e. toothed whales) approximately 32 million years ago (McGowen et al., 2014). It is the sole extant member of the *Physeter* genus and is most closely related to pygmy (*Kogia breviceps*) and dwarf (*Kogia sima*) sperm whales (McGowen et al., 2014). The scientific name of the species, *Physeter macrocephalus*, roughly translates to "big-headed blowhole" in Greek, and aptly highlights sperm whales' most notable anatomical feature: their massive, boxy heads. These heads, which comprise 25–33% of an adult sperm whale's total body length (~16 m for males, ~11 m for females) (Whitehead, 2018), house what could arguably be called the "superpower" of the species: the spermaceti organ.

The spermaceti organ contains a liquid mixture of fats and waxes, called spermaceti (Clarke, 1978). It enables sperm whales to create highly directional, powerful clicks, which they use for foraging and communicating (Cranford, 1999). Unfortunately for the whales, spermaceti made them a major target for commercial whalers, as it could be used in many human products, including lamps, lubricants, and candles[2] (Irwin, 2012; Rubenstein, 1963). It was also spermaceti that gave sperm whales their common name, as early whalers mistook the substance for sperm when processing dead whales (paradoxically, even when the dead whales were female).

Hundreds of thousands of sperm whales were killed by whalers in the 19th and 20th centuries (Cressey, 2015; Whitehead, 2002; Whitehead et al., 2021). It was not until 1982 that the International Whaling Commission announced a moratorium on commercial whaling, which came into effect during the 1985/1986 seasons[3]. It has been estimated that the number of sperm whales lost to whaling has resulted in two million extra tons of carbon remaining in the atmosphere each year—an astonishing 70 million tons since the moratorium went into effect (Lavery et al., 2010). The current sperm whale population is thought to be roughly a third of the pre-whaling population (Whitehead, 2002).

Sperm whales were whaled worldwide because they have a cosmopolitan distribution, but this distribution is structured by sex and age (Jaquet, 1996). Female and immature sperm whales generally inhabit areas with sea surface temperatures warmer than 15°C, which means they are largely found between the 40° latitudinal parallels (Rice, 1989). In contrast, male sperm whales disperse from their natal groups sometime between four and 21 years of age (before they become sexually mature) and spend more time in polar waters as they grow older (Best, 1979). Sexually mature males return to

---

[2] In late 19th century England, the "candlepower" was a standard unit of light based on the amount of light produced by a pure spermaceti candle (Treese, 2018).
[3] The International Whaling Commission 34th annual meeting minutes (1982) are archived at: https://archive.iwc.int/pages/view.php?ref=424&k=.

warmer waters to breed, but otherwise live largely separate lives from female and immature sperm whales (Whitehead & Weilgart, 2000).

Due in no small part to ease of access, much sperm whale research over the past 50 years has focused on female and immature whales (but for research on males, see Jaquet et al., 2001; Frantzis & Alexiadou, 2008; Steiner et al., 2012; Straley et al., 2014; Guerra et al., 2017; Kobayashi et al., 2020). From this work, we have learned that sperm whale societies are hierarchically structured. Related and unrelated individuals live in stable, matrilineal 'social units' (Christal et al., 1998; Whitehead et al., 1991). Social units will associate for hours, days, or weeks at a time, forming 'groups' (Christal et al., 1998; Whitehead et al., 1991). In social situations, sperm whales communicate using patterned series of clicks, called codas (Watkins & Schevill, 1977). By varying the number and spacing (i.e. inter-click interval; ICIs) of clicks, whales can produce coda types with different rhythms and tempos (Weilgart & Whitehead, 1997). Social units show clear preferences (which are socially transmitted among unit members) for certain coda types and will only associate with other units that have similar preferences (i.e. that share their coda dialect) (Cantor et al., 2015; Gero, Whitehead, et al., 2016). These preferential interactions give rise to a cultural level of population structure known as the 'vocal clan' (Rendell & Whitehead, 2003b).

Culture, defined as information or behavior that is shared within a community and acquired from conspecifics through social learning (Whitehead & Rendell, 2014), can significantly structure populations and has been called a second inheritance system (Whiten, 2017). It may have even helped sperm whales mitigate some of the effects of commercial whaling (Whitehead et al., 2021). Vocal clans were first described in the eastern tropical Pacific in 2003 (Rendell & Whitehead, 2003b), but hints of their existence are present in older studies (e.g. Weilgart & Whitehead, 1997; Christal et al., 1998). Sympatric vocal clans have been documented off Brazil (Amorim et al., 2020), Mauritius (Huijser et al., 2020), and Japan (Amano et al., 2014) as well, while the Mediterranean Sea is home to a single clan (Rendell & Frantzis, 2016).

From long-term research projects in the Galápagos Islands and Dominica, we have learned that different vocal clans not only vocalize differently (Gero, Bøttcher, et al., 2016; Rendell & Whitehead, 2003b), but also forage (Marcoux, Whitehead, et al., 2007; Whitehead & Rendell, 2004), move (Whitehead & Rendell, 2004), dive (Cantor & Whitehead, 2015), associate (Cantor & Whitehead, 2015), and distribute (Vachon et al., submitted; Eguiguren et al., 2019) differently, mirroring human cultural differences in accent, music, cuisine, traditions, and distribution. These cultural differences can result in fitness variation among different sperm whale clans, including differences in reproductive (Marcoux, Rendell, et al., 2007) and foraging success (Whitehead & Rendell, 2004).

Despite the many advances in our understanding of sperm whale society and culture over the past few decades, there are still many unknowns, particularly regarding vocal clans. How many clans exist worldwide? Are sympatric clans the norm or the exception? Are clan dialects stable over time? Do codas act as symbolic markers of clan identity? How do clan coda preferences vary, and which structural features of codas are important? These are questions I revisited throughout my Ph.D., as I sought to better understand how time and space impact sperm whale social communication.

## 1.6 – Thesis overview and objectives

The overarching goal of my thesis is to investigate how sperm whale dialects vary over space and time, and to determine the implications of these spatiotemporal dynamics on cultural identity and evolution in clans. Throughout my degree, I drew inspiration from various research areas and topics, including biomusicology, fuzzy clustering, linguistics, ethnogeography, and cultural evolution. My work heavily relied on acoustic recordings collected around the world over the past 50 years and highlights the insights that can be gained from collaborative, long-term, and large-scale acoustic datasets.

In the second chapter, I survey the cetacean literature for evidence of rhythm in mysticete (i.e. baleen whale) and odontocete vocalizations and use an existing definitional framework from Ravignani et al. (2014) to summarize my findings. I show that rhythm (particularly isochronous rhythm) is common in cetacean vocalizations but may be used in different behavioral contexts by mysticetes and odontocetes. Using a comparative lens, I discuss the evolutionary advantages of vocal rhythm and theorize about what role it may play in different cetacean vocal displays.

In the third chapter, I introduce a new method—the 'identity call method', or IDcall—for detecting putative biological structure in acoustic datasets. IDcall classifies calls into types using contaminated mixture models; hierarchically clusters repertoires of calls based on similarities in call type usage; and produces a dendrogram with 'identity clades' and the 'identity calls' that best characterize each clade. The method also provides information on usage of 'non-identity calls' (i.e. call types that are used by more than one identity clade and are thus not unique to any one clade). The detected identity clades match up well with known sperm whale vocal clans, grey-breasted wood-wren (*Henicorhina leucophrys*) subspecies, and Australian field cricket (*Teleogryllus*) species, showing that the method can be applied to different taxa, types of vocalizations, and levels of biological structure. This chapter was published in the journal *Methods in Ecology and Evolution* in 2021 and was the cover of the September issue.

Within a species, dialect variation can result from both spatial and temporal factors. My fourth and fifth chapters address each of these possibilities for sperm whales, respectively. Both chapters use IDcall to determine how many sperm whale clans are present in the corresponding coda datasets and to characterize each clan's coda dialect.

In the fourth chapter, I analyze over 23,000 codas from 23 Pacific locations to characterize sperm whale culture across an entire ocean basin. I document seven sperm whale clans in the Pacific and show that clan distributions can vary dramatically, with some clans restricted to small areas and others spanning the ocean. Within clans, I show that usage of identity codas is more stable over geographic space than usage of non-identity codas. Between clans, I show that similarity in identity coda usage decreases with increasing spatial overlap, whereas non-identity coda usage does not change. These results provide empirical evidence that identity codas function as symbolic markers of sperm whale cultural identity, like ethnic markers in humans.

In the fifth chapter, I investigate the fine-scale structural stability of well-sampled sperm whale coda types over decadal and multi-decadal timespans. This chapter focuses on the coda dialects of five clans from three geographic regions: the Regular and Plus-One clans from the Galápagos Islands (Rendell & Whitehead, 2003b); the eastern Caribbean 1 (EC1) and eastern Caribbean 2 (EC2) clans from the Lesser Antilles (Gero, Bøttcher, et al., 2016); and the Mediterranean clan from the Balearic Islands (Rendell &

Frantzis, 2016). While within-type coda similarity generally decreases over time across clans, I show that fine-scale levels of temporal stability differ by coda type. For types that have evolved, the magnitude and direction of change can vary as well, with some coda type durations/ICI ratios increasing while others decrease. The results suggest that codas are susceptible to cultural drift and are more graded than has previously been recognized.

Finally, in chapter six, I summarize and contextualize my thesis chapters, recap the past and future of sperm whale research, and consider the broader implications of my results. This chapter is followed by four appendices containing supplementary material for chapters 2–5.

# CHAPTER 2

# RHYTHM IN CETACEAN VOCALIZATIONS[4]

> "I have begun to internalize the rhythms, including the rhythms by which the whales, at least, keep their promises."
>
> —Carl Safina

## 2.1 – Abstract

Vocal rhythms can significantly augment sociality and behavior, but studies have generally been limited in taxonomic scope. Researchers are calling for a comparative, cross-species approach to better understand the adaptive value and evolutionary history of vocal rhythm production. This review synthesizes the current state of knowledge about rhythm in vocalizations for one promising taxa—cetaceans—and interprets it under a comparative lens. We[5] define rhythm and summarize a pre-existing definitional framework that can be used to guide vocal rhythm research. Using this framework, we describe examples of vocal rhythm in 18 cetacean species (7 mysticetes, 11 odontocetes). Cross-species trends suggest that while rhythm (particularly isochronous rhythm) is prevalent in cetacean vocalizations, it is primarily evident in song in mysticetes but used across diverse vocalization types in odontocetes. This disparity may stem from differences in mysticete and odontocete anatomies, foraging strategies, and social structures. Inspired by some of the limitations encountered in the cetacean literature, we suggest steps bioacousticians can take to better situate their work within the broader field, as well as some exciting avenues for future research. By explicitly quantifying vocal rhythm production at various scales and in diverse species, and by being more intentional when describing vocal rhythm production, cetacean researchers specifically, and

---

[4] This chapter is in collaboration with Hal Whitehead. TAH conceived the study, conducted the literature review, and drafted the manuscript. HW provided input at all stages.

[5] Chapters 2–5 were done in collaboration with various researchers, who are acknowledged at the beginning of each chapter. I use collective pronouns throughout chapters 2–5 to reflect this, but I have made the most substantial contribution to each chapter in this thesis.

bioacousticians more broadly, can contribute to a rapidly growing and cross-species body of work.

## 2.2 – Introduction

We live in a rhythmic world. From large-scale geophysical rhythms, like seasons and tides, to fine-scale anatomical and neurological rhythms, like heartbeats and brainwaves, rhythms abound across all aspects of life (Fitch, 2012). In humans (*Homo sapiens*), rhythmic perceptual abilities emerge as early as two months of age and continue to develop throughout life (Demany et al., 1977; Drake et al., 2000). The impacts of these abilities are far-reaching, as they allow us to perceive, interpret, and create rhythm in music (Bispham, 2006), language (Langus et al., 2017), and dance (Bresnahan, 2019). In the acoustic domain, rhythm can markedly improve our ability to detect (Rimmele et al., 2011), react to (Rimmele et al., 2011), and compare (Jones et al., 2002) signals. Rhythm also facilitates future-directed attending by allowing attention to be optimally targeted to specific points in time (Large & Jones, 1999). This, in turn, enables joint attention and behavioral synchronization, which have allowed humans to fundamentally alter the planet (McNeil, 1997). Despite the impressive ability of rhythm to facilitate or enhance various aspects of human sociality and behavior, there are still more questions than answers regarding the evolution of vocal rhythm production (Fitch, 2012). Many of the gaps in our knowledge could be filled by embracing a comparative approach and looking at rhythmic abilities in non-human animals (Bispham, 2006; Fitch, 2013; Kershenbaum, Blumstein, et al., 2016; Merker et al., 2009; Ravignani, Dalla Bella, et al., 2019).

Comparative approaches have been illuminating in other acoustic research areas (e.g. song production, Garland & McGregor, 2020; vocal learning, Lattenkamp & Vernes, 2018) and can help discern the evolutionary trajectory, selective pressures, underlying mechanisms, and functions of traits across species. In research on vocal rhythm production, the bulk of comparative work has been done on non-human primates (e.g. Ghazanfar, 2013) and passerine birds (e.g. Norton & Scharff, 2016). More recently, the field has broadened to include bats (e.g. Burchardt et al., 2019) and pinnipeds (e.g. Ravignani, 2019b), but the diversity of taxa for which vocal rhythm production has been explicitly considered remains low. Here, we argue that cetaceans (i.e. whales, dolphins, and porpoises) represent a fruitful next research frontier.

Similar to pinnipeds (Ravignani et al., 2016), cetaceans may shed light on the origins of human vocal rhythm since they are evolutionarily closer to humans than passerine birds and share more vocal capabilities (e.g. vocal learning; Tyack & Sayigh, 1997) with humans than non-human primates. Several cetacean species also exhibit culture (reviewed in Whitehead & Rendell, 2014), which has known interactions with rhythm production, transmission, and perception in humans (e.g. Hannon et al., 2012; Jacoby & McDermott, 2017; Ravignani et al., 2018). For example, categorical rhythms (those in which "temporal intervals between note onsets are distributed categorically rather than uniformly"; De Gregorio et al., 2021) are thought to promote cultural transmission of learned vocalizations in humans and birds (Roeske et al., 2020), and could be operating in similar ways in cetaceans with culturally-transmitted dialects, like sperm whales (*Physeter macrocephalus;* Cantor et al., 2015).

Studying cetacean vocal rhythm production will also help us better understand how these animals utilize sound in a complex acoustic environment, where other senses are comparatively limited. Within cetaceans, mysticetes (i.e. baleen whales) and odontocetes (i.e. toothed whales) differ significantly in anatomies, foraging strategies, and social structures, which could impact patterns in vocal rhythm production. Learning more about rhythm in cetacean vocalizations[6] can also inform hypotheses and theories in research areas beyond just vocal rhythm, including biomusicology (Fitch, 2015) and entrainment (Wilson & Cook, 2016). Ultimately, many of the ways in which vocal rhythm augments sociality in humans and other taxa could be at play in cetaceans.

Here, we summarize the current state of knowledge on the production of periodic rhythms in cetacean vocalizations. Our approach illustrates one route researchers can take to induct new species into a comparative approach to vocal rhythm production, and we encourage others to join the fray: broadening the diversity of species studied allows us to broaden the diversity of questions asked. We begin by defining rhythm (see Table 2.1 for a glossary of key terms), followed by a brief summary of a conceptual framework that

---

[6] Throughout this review, we use the word 'vocalization' to mean any acoustic call produced by an animal, regardless of whether the call is produced via the vocal tract/cords (as some definitions of 'vocalization' require) or via other anatomical features (e.g. the phonic lips in odontocetes).

can help scaffold and standardize vocal rhythm research (Ravignani et al., 2014). Guided by this framework, we then summarize what is currently known about rhythm in cetacean vocalizations (i.e. rhythm production) and interpret the trends under a comparative lens. We conclude by discussing steps bioacousticians can take to situate their work within a broader comparative framework for vocal rhythms and highlighting future avenues of research. To our knowledge, this review is the first infraorder-level assessment of vocal rhythm production to date and the most systematic application of any conceptual framework of rhythm on a large taxonomic scale.

*Table 2.1* – Glossary of key terms.

| **Term** | **Definition** |
|---|---|
| Aperiodic | "Non-repeating"[7] |
| Call | A single continuous vocalization flanked by silences |
| Cetaceans | Species comprising the Cetacea infraorder; includes whales, dolphins, and porpoises |
| Chorus tree | Definitional framework used to describe vocalizations from groups of animals[7] |
| Complex interval ratios | Ratios of relatively large integers[5] (e.g. $1 - {}^{15}/_{17} - {}^{19}/_{28} - 1 - {}^{15}/_{17} - {}^{19}/_{28} -$ 1...) |
| Expressivity | "Capacity of a signaling system to convey different meanings"[8] |
| Heterochronous/ heterochrony | Periodic events (e.g. sounds) are separated by intervals of more than one duration[7] |
| Interval | The silent period that separates events (e.g. sounds)[7] |
| Isochronous/ isochrony | Periodic events (e.g. sounds) are separated by an interval of a single duration; metronomic[7] |
| Mysticetes | Baleen whales |
| Odontocetes | Toothed whales |
| Periodic | "Regularly repeating"[7] |
| Rhythm | "pattern of time intervals demarcating a sequence of stimulus events"[9] |
| Quasiperiodic | "Nearly but not perfectly regular"[7] |
| Simple interval ratios | "Ratios of relatively small integers"[1] (e.g. 1 – ¼ – ¼ – 1 – ¼ – ¼ – 1...) |
| Solo tree | Definitional framework used to describe vocalizations from single animals[7] |
| Song | Vocalizations known or hypothesized to have a role in courtship/sexual advertising by males |
| Tempo | The rate of events (e.g. sounds) over time |

---

[7] Definitions from Ravignani et al. (2014).
[8] Definitions from Ravignani & Madison (2017).
[9] Definitions from Leow & Grahn (2014).

### *2.2.1 – Defining rhythm*

A critical first step in assessing rhythm in cetacean vocalizations is to define rhythm because, to borrow from Schuppli and van Schaik (2019), the way we define rhythm has significant implications for where we find it. The *New Grove Dictionary of Music and Musicians* defines rhythm as “a series of durations or timespans which form groups” (London, 2001). Patel (2008) and Norton and Scharff (2016) opt for rhythm as the “systematic patterning of sound in terms of timing, accent, and grouping.” Some definitions invoke specific timescales ("temporal structure at a second-millisecond time scale"; Ravignani, 2019a) while others are quite broad ("a regular repeated pattern"; Spierings & ten Cate, 2016). The definition of rhythm we will adopt is a “pattern of time intervals demarcating a sequence of stimulus events” (Leow & Grahn, 2014). We like this definition for several reasons: (1) it requires some degree of repetition (via the word ‘pattern’[10]) and at least two intervals (via the plurality of ‘intervals’) for a sequence to be classified as rhythmic (Bouwer et al., 2021); (2) it remains flexible by not focusing on specific event (e.g. call) types or specifying an explicit timescale, and; (3) it is broad enough to be applied across species. Given the typical duration of cetacean vocal displays and recordings, however, the timescales included here are typically on the order of seconds or minutes for mysticetes and milliseconds or seconds for odontocetes.

### *2.2.2 – Studying rhythm*

Methods for quantifying rhythm in vocalizations have been extensively reviewed (Kershenbaum et al., 2016; Ravignani & Norton, 2017) and compared (Burchardt & Knörnschild, 2020; Schneider & Mercado III, 2019) elsewhere and will not be reiterated here. Vocal rhythm analyses can be guided by and centered within conceptual frameworks (e.g. Anichini et al., 2020; Kershenbaum et al., 2016; Ravignani, 2019a). Here, we employ a definitional framework proposed by Ravignani et al. (2014) to describe rhythm in cetacean vocalizations. This framework uses descriptive hierarchies to categorize temporal patterns and is appealing because it: (1) addresses the pervasive issue of inconsistent terminology in rhythm analyses, thereby facilitating a more comparative

---

[10] The Cambridge Dictionary defines pattern as “any **regularly repeated** arrangement” (emphasis our own) (definition retrieved September 18, 2021, from https://dictionary.cambridge.org/dictionary/english/pattern).

approach; (2) emphasizes that rhythm is a composite (i.e. multi-component) system (Fitch, 2012) by using hierarchies of rhythmic structure to describe vocalizations; and (3) focuses on categorizing rhythmic patterns, not on the mechanisms that enable their production (Ravignani et al., 2014). This last feature is especially key for describing cetacean vocal rhythms, given that our ability to deduce mechanisms is often limited by the location and life history of our study subjects.

Under this framework, vocalizations produced by single individuals (i.e. the "solo tree") are first classified as *periodic* (i.e. regularly repeating) or *aperiodic* (i.e. non-repeating) (Figure 2.1). All periodic signals are rhythmic under our definition and are further characterized as either *isochronous* (i.e. sounds are separated by a silent interval of a single duration; metronomic) or *heterochronous* (i.e. sounds are separated by silent intervals of more than one duration; intervals are related by *simple* (e.g. 1 – ⅓ – ⅓ – 1 – ⅓ – ⅓ – 1…) or *complex* (e.g. 1 – $^{7}/_{16}$ – $^{11}/_{15}$ – 1 – $^{7}/_{16}$ – $^{11}/_{15}$ – 1…) *ratios*). Although not repeated, aperiodic sound intervals can also be related by simple or complex ratios. Whether a signal is considered periodic or aperiodic can heavily depend on the time scale analyzed (Ravignani et al., 2014). For example, a given pattern might appear aperiodic over seconds, periodic over minutes, and aperiodic over hours. Additionally, few phenomena are "perfectly periodic"—even atomic clocks drift—and how close a signal is to perfectly periodic can fluctuate throughout the signal duration (Fitch, 2012; Ravignani & Madison, 2017). For simplicity, we use the terms isochronous/isochrony and heterochronous/heterochrony[11] throughout this review to mean signals that exhibit local or quasiperiodic (i.e. nearly but not perfectly regular; Ravignani et al., 2014) behavior at any time scale. Using this hierarchy, research on vocalizations can be distilled down to succinct yet informative summaries of rhythmic structure.

---

[11] Note that the term 'heterochrony' is polysemic. In evolutionary developmental biology, it refers to a change in the timing of developmental processes in an organism compared to its ancestor. In rhythm analysis, it refers to when periodic events, such as sounds, are separated by silent intervals of more than one duration (see Table 2.1).

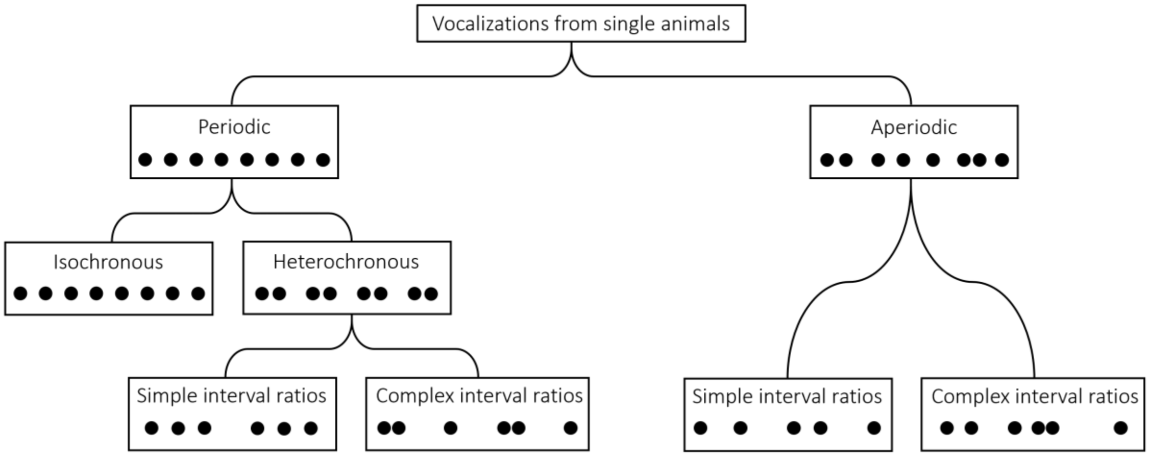


*Figure 2.1* – Definitional framework for characterizing rhythm in vocalizations from single animals. This is an adaptation of the "solo tree" from Ravignani et al. (2014). Vocalizations can be described by starting at the top of the tree and choosing the option that best characterizes the vocalization at each bifurcation. Visual examples are shown beneath each descriptor and depict the timing of a sequence of vocalizations (black dots), with spaces between the dots representing inter-vocalization (i.e. silent) intervals.

## 2.3 – Methods

Studies on vocal rhythm production in cetaceans were identified using a combination of a scoping literature review of four abstract and citation databases (Biological Abstracts, Web of Science Core Collection, ProQuest, and Scopus) and extensive “hand searches” of Google Scholar (Method S2.1 in Appendix A). Results from relevant studies were synthesized into descriptions of rhythmic behavior for mysticete and odontocete species using a definitional framework (Ravignani et al., 2014). Evidence for vocal rhythm was categorized as quantitative, descriptive (i.e. explicitly described as rhythmic in the text), visual (i.e. inferred from spectrograms, chronograms, figures, etc.), or unpublished[12]. For the purposes of this review, only quantitative evidence was considered definitive, while descriptive, visual, and unpublished evidence were considered preliminary. Only definitive evidence is discussed in the main text and Tables (but see the ‘Future research directions’ section and Tables S2.1/S2.2 for additional information on the preliminary evidence).

---

[12] Unpublished data were acquired after reaching out to corresponding authors on papers found in our scoping review/hand searches.

Many studies did not specifically aim to quantify rhythm in cetacean vocalizations but provided metrics that we could retroactively use to assess rhythm. By far the most common metric we used was the coefficient of variation (CV) of inter-unit intervals, which is calculated by dividing the standard deviation of the inter-unit intervals by the mean. The CV is a useful but limited measurement for making rhythmic inferences, as low CVs indicate isochrony but high CVs could indicate heterochrony or aperiodicity. Furthermore, no ‘thresholds’ exist for how low a CV must be to indicate isochrony (Burchardt et al., 2021)—nor, in our opinion, should they, given that isochrony is likely a continuum rather than a binary feature of vocalizations and will undoubtedly vary depending on the temporal perception and resolution abilities of different species. As such, we imposed no absolute thresholds here, but note that the closer a CV is to 0, the stronger the evidence that a given sequence is isochronous. All CVs can be found in Tables S2.1/S2.2.

As the goal of this review was to provide an overview of the presence of and trends in cetacean vocal rhythm, the evidence presented here is in-depth but should not be considered exhaustive. Many of the papers we read described aperiodic cetacean vocalizations (e.g. northern bottlenose whale *(Hyperoodon ampullatus)* surface clicks, Hooker & Whitehead, 2002; blue whale *(Balaenoptera musculus)* D calls, Oleson et al., 2007), which are not covered beyond this point.

## 2.4 – Results

### *2.4.1 – Rhythm in mysticete vocalizations*

There is definitive evidence of rhythm in vocalizations for a diversity of mysticete species, including blue, bowhead *(Balaena mysticetus)*, fin *(Balaenoptera physalus)*, humpback *(Megaptera novaeangliae)*, minke (*Balaenoptera acutorostrata;* dwarf subspecies), North Pacific right *(Eubalaena japonica)*, and Omura’s *(Balaenoptera omurai)* whales (Tables 2.2/S2.1). There was clear evidence of isochrony in calls from various species (e.g. Omura's whale 15-50 Hz amplitude-modulated calls; Cerchio et al., 2015), with calls defined here as single vocalizations. All the definitive examples of rhythm were in vocalizations known or hypothesized to have a role in courtship/sexual advertising by males (i.e. song).

*Table 2.2* – Definitive examples of isochronous and heterochronous rhythm in mysticete vocalizations. For each vocalization, the known or hypothesized behavioral context is taken from the source(s) in the final column. Summaries of vocal rhythm are based on quantitative evidence only. See Table S2.1 for a more detailed version of this table, which includes descriptions of vocalization components, additional preliminary examples of rhythm in mysticete vocalizations, additional references, and CVs (when possible).

| **Common name** | **Vocalization** | **Context** | **Summary of vocal rhythm** | **Rhythm summary**[13] | **References** |
|---|---|---|---|---|---|
| Blue whale | Song | Courtship | Heterochronous calls in non-contiguous A-B phrases; isochronous phrases in sequences; isochronous sequences in songs | I, H, * | (Mellinger and Clark, 2003; Oleson et al. 2007; Stafford et al. 2001) |
| Bowhead whale | Song | Courtship | Isochronous songs in song bouts | I | (Delarue, Laurinolli, et al., 2009; Stafford et al., 2008) |
| Fin whale | Song | Courtship | Isochronous and heterochronous 20 Hz pulses in songs | I, H | (Delarue, Todd, et al., 2009; Pereira et al., 2020; Širović et al., 2017) |
| Humpback whale | Song | Courtship | Isochronous and heterochronous units in songs | I, H | (Handel et al., 2009; Schneider & Mercado III, 2019) |
| Minke whale (dwarf subspecies) | Song | Courtship, socializing, spacing | Isochronous star wars vocalizations in slow and fast songs; heterochronous star wars vocalizations in rapid-clustered songs | I, H | (Gedamke, 2004; Gedamke et al., 2001) |
| North Pacific right whale | Song | Courtship | Isochronous and heterochronous calls in song phrases; isochronous phrases in songs | I, H, * | (Crance et al., 2019) |
| Omura's whale | Song | Courtship | Isochronous 15-50 Hz amplitude-modulated calls in songs | I | (Cerchio et al. 2015) |

---

[13] I = evidence for isochronous rhythm; H = evidence for heterochronous rhythm; * = evidence for rhythm at multiple levels in hierarchical vocalization (see Table S2.1 for more details)

The stark prevalence of song-based examples of rhythm may reflect a research bias towards studying song (indeed, some definitions of song include rhythm as a diagnostic feature, e.g. Crance et al., 2019), but also suggests that rhythm may be a ubiquitous feature of mysticete song. All the song/courtship-related vocalizations we examined have evidence of isochrony. While the complexity of the calls themselves varies (e.g. simple fin whale 20 Hz pulses vs. complex dwarf minke whale star-wars vocalizations; Delarue, Todd, et al., 2009; Gedamke et al., 2001), the inter-call interval is often quite regular and stereotyped within songs. For species with hierarchically organized songs, like North Pacific right whales, there is evidence of isochrony at multiple levels of the hierarchy (e.g. inter-unit and inter-phrase intervals; Crance et al., 2019).

Heterochrony seems to be somewhat rarer[14] than isochrony in mysticetes, but was observed in the songs of five species (blue, fin, humpback, dwarf minke, and North Pacific right whales) and usually manifested as two (e.g. dwarf minke whale rapid-clustered songs, Gedamke, 2004; blue whale A-B song, Oleson et al., 2007; fin whale doublet song, Širović et al. 2017) or three (e.g. North Pacific right whale GS2-TP and GS4-DG songs; Crance et al., 2019) inter-call intervals within single songs. Some of the earliest spectrograms of humpback whale songs (e.g. Guinee & Payne, 1988; Payne & McVay, 1971) suggest periods of heterochrony, which was confirmed by recent quantitative work that found that humpback songs have heterochronic, isochronic, quasiperiodic, shifting, and aperiodic segments (Schneider & Mercado III, 2019). Despite decades of extensive work on humpback whales, however, very few studies explicitly quantify silent intervals within and between songs, which limits rhythmic inferences. In a similar vein, quantification of silent intervals at the phrase level appears to be rarer than at the call level in species with hierarchical song, like humpback, bowhead, and North Pacific right whales. Most studies to date have focused on spectral and temporal features of the sounds themselves, despite cross-species evidence (e.g. Gerhard, 2003; Margoliash, 1983; Williams & Staples, 1992) that silences can be just as or more important than sounds in rhythm production and perception.

---

[14] Throughout this review, when we say something is rare, we acknowledge that it may not actually be a rare phenomenon, but that it has been rarely documented in the scientific literature.

### *2.4.2 – Rhythm in odontocete vocalizations*

Vocalizations from various odontocete species—Atlantic spotted dolphins *(Stenella frontalis)*, beluga whales *(Delphinapterus leucas)*, Blainville's beaked whales *(Mesoplodon densirostris)*, common bottlenose dolphins *(Tursiops truncatus)*, Cuvier's beaked whales *(Ziphius cavirostris)*, Indo-Pacific bottlenose dolphins *(Tursiops aduncus)*, orcas *(Orcinus orca)*, long-finned pilot whales *(Globicephala melas)*, northern bottlenose whales, northern right whale dolphins *(Lissodelphis borealis)*, and sperm whales—have evidence of rhythm (Tables 2.3/S2.2). As in mysticetes, we found examples of isochronous (e.g. sperm whale surface clicks, Jaquet et al., 2001) and heterochronous (e.g. common bottlenose dolphin synchronized whistle/buzz bouts, Herzing, 2015) rhythm in vocalizations related to courtship, but rhythm was also abundant in foraging, social, and aggressive vocalizations.

*Table 2.3* – Definitive examples of isochronous and heterochronous rhythm in odontocete vocalizations. For each vocalization, the known or hypothesized behavioral context is taken from the source(s) in the final column. Summaries of vocal rhythm are based on quantitative evidence only. See Table S2.2 for a more detailed version of this table, which includes descriptions of vocalization components, additional preliminary examples of rhythm in odontocete vocalizations, additional references, and CVs (when possible).

| **Common name** | **Vocalization** | **Context** | **Summary of vocal rhythm** | **Rhythm summary**[15] | **References** |
|---|---|---|---|---|---|
| Atlantic spotted dolphin | Screams | Aggression | Isochronous screams in series | I | (Herzing, 2015) |
| | Squawks | Aggression | Isochronous squawks in series | I | (Herzing, 2015) |
| Beluga whale | Echolocation | Foraging, sensing | Isochronous clicks in series | I | (O. Le Bot et al., 2015) |
| Blainville's beaked whale | Echolocation | Foraging, sensing | Isochronous clicks in series | I | (Johnson et al., 2006) |
| Common bottlenose dolphin | Disconnected multi-loop signature whistles | Socializing | Isochronous loops in disconnected multi-loop signature whistles | I | (Esch et al., 2009) |
| | Bray/buzz bout series | Aggression, courtship | Isochronous bray/buzz bouts in series | I | (Herzing, 2015) |
| | Buzz bout series | Aggression, courtship | Isochronous buzzes in series | I | (Herzing, 2015) |
| | Whistle/buzz bout series | Aggression, courtship | Heterochronous whistles and buzzes in series; isochronous bouts in series | I, H, * | (Herzing, 2015) |

*Continued on next page*

---

[15] I = evidence for isochronous rhythm; H = evidence for heterochronous rhythm; * = evidence for rhythm at multiple levels in hierarchical vocalization (see Table S2.2 for more details)

| **Common name** | **Vocalization** | **Context** | **Summary of vocal rhythm** | **Rhythm summary** | **References** |
|---|---|---|---|---|---|
| Cuvier's beaked whale | Echolocation | Foraging, sensing | Isochronous clicks in series | I | (Zimmer et al., 2005) |
| Indo-Pacific bottlenose dolphin | Disconnected multi-loop signature whistles | Socializing | Isochronous loops in disconnected multi-loop signature whistles | I | (Gridley, 2010) |
| | Pop train | Courtship | Isochronous pops in pop trains | I | (Moore et al., 2020) |
| Long-finned pilot whale | Repeated call sequences | Socializing | Isochronous calls in repeated call sequences | I | (Vargas, 2017) |
| Northern bottlenose whale | Echolocation | Foraging, sensing | Isochronous clicks in series | I | (Hooker & Whitehead, 2002) |
| Northern right whale dolphin | Burst pulse series | Unknown | Isochronous clicks in burst pulse units; heterochronous burst pulse units in series | I, H, * | (Rankin et al., 2007) |
| Orca | Discrete call series | Socializing, social travelling | Isochronous discrete calls in series | I | (Miller et al., 2004) |
| | Ultrasonic whistle series | Unknown | Isochronous ultrasonic whistles in series | I | (Simonis et al., 2012) |

*Continued on next page*

| **Common name** | **Vocalization** | **Context** | **Summary of vocal rhythm** | **Rhythm summary** | **References** |
|---|---|---|---|---|---|
| Sperm whale | Codas | Socializing | Isochronous and heterochronous clicks in codas; isochronous codas in bouts | I, H, * | (Amorim et al. 2020; Cantor et al. 2016; Gero et al. 2016; Huijser et al. 2020; Schulz et al. 2008; Weilgart and Whitehead 1997) |
| | Echolocation | Foraging, sensing | Isochronous clicks in series | I | (Burchardt & Knörnschild, 2020; Tønnesen et al., 2020) |
| | Surface clicks | Courtship | Isochronous clicks in series | I | (Jaquet et al., 2001) |

Echolocation click trains from several species (Table 2.3) often have periods with isochronous clicks, but the inter-click interval can vary based on where an animal is in their dive cycle ( Linnenschmidt et al., 2013), distance to prey (Doh et al., 2018), prey behavior (Guerra et al., 2017), habitat (Yamamoto et al., 2016), and time of day (Yamamoto et al., 2016). The inter-click intervals of neighboring clicks generally smoothly change (what Schneider and Mercado III (2019) would call "shifting" and musicians might call "accelerando" or "ritardando") during periods of click acceleration and deceleration (e.g. Backus & Schevill, 1966). These results mirror those from another echolocating species, the greater sac-winged bat (*Saccopteryx bilineata;* Burchardt et al., 2019). Orca ultrasonic whistles may be used in foraging and can exhibit isochrony as well (Filatova et al., 2012; Simonis et al., 2012).

Isochronous patterns were also seen within and between calls with known or presumed social/affiliative functions in various species. For example, the spacing between loops (i.e. repeated elements) within some multi-loop signature whistles (which signify individual identity) is isochronous for both common (Esch et al., 2009) and Indo-Pacific (Gridley, 2010) bottlenose dolphins. In long-finned pilot whale repeated call sequences (Vargas, 2017; Zwamborn & Whitehead, 2017) and sperm whale coda bouts (Schulz et al., 2008), the vocalizations are isochronously spaced. Herzing (2015) also documented isochronous rhythm in vocalizations heard during aggressive interactions between Atlantic spotted dolphins and common bottlenose dolphins in the Bahamas.

Compared to isochronous rhythm, heterochronous rhythm was rarely quantified in odontocete vocalizations. Definitive evidence of heterochrony exists for at least three species: common bottlenose dolphins, northern right whale dolphins, and sperm whales. Eight different types of burst pulse series recorded from northern right whale dolphins were all heterochronous (Rankin et al., 2007). Sperm whale codas can be isochronous (e.g. the '5R' coda, Rendell & Whitehead, 2003) and heterochronous (e.g. the '4+1' coda, Rendell & Whitehead, 2003) and some vocal clans of sperm whales show clear preferences for overarching rhythmic patterns in their codas; for example, the ‘Regular’ clan primarily makes isochronous codas whereas the ‘Plus-One’ and 'Four-Plus’ clans primarily make heterochronous codas (Rendell & Whitehead, 2003b). Similar to North

Pacific right whale song, sperm whale codas highlight how rhythm can be present at multiple levels in hierarchical cetacean vocal displays (Schulz et al., 2008).

## 2.5 – Discussion

Taken together, the trends in mysticete and odontocete vocalizations lend themselves to four general conclusions:

1. Rhythm is a common feature of cetacean vocalizations.
2. Isochrony is more common (and/or better documented) than heterochrony in cetacean vocalizations.
3. Rhythm in mysticete vocalizations is evident in song.
4. Rhythm in odontocete vocalizations occurs across diverse vocalization types.

Below, we discuss possible explanations for each conclusion under a comparative lens.

### *2.5.1 – Commonality of rhythm*

As Tables 2.2/2.3/S2.1/S2.2 show, there are many examples of definitive and probable rhythm in cetacean vocalizations. To understand why, it is helpful to consider how rhythm functions in the vocalizations of other species (this section) and how specific aspects of cetacean ecology, anatomy, and sociology might foster vocal rhythm production (subsequent sections). Comparative research shows that rhythm can play nontrivial roles in memory, attention, and synchronization across species.

In humans, rhythm provides mnemonic benefits in perception and memory tasks, and these benefits can be cross modal (Hickey et al., 2020; Jones & Ward, 2019). Isochronous rhythm can dynamically modulate episodic memory, improving recognition (e.g. Jones & Ward, 2019), reaction time (e.g. Hickey et al., 2020), and incidental learning (e.g. Thavabalasingam et al., 2016). The memory benefits conferred by rhythm are not restricted to humans; young zebra finches (*Taeniopygia guttata*) learn new song sequences more accurately (in terms of similarity to a tutor song) when the songs are isochronous vs. aperiodic (Hyland Bruno, 2017). Rhythm may provide similar mnemonic benefits in cetaceans, enabling diverse vocal repertoires and complex vocal displays to be learned and remembered (Handel et al., 2009). How exactly can rhythm improve memory? Modulation of attention is key.

Rhythm is powerful because it can guide temporal allocation of attention, allowing attention to be optimally (and economically) targeted at expected time points and facilitating future-directed attending (Bermeitinger & Frings, 2015; Bispham, 2006; Huron, 2008; Large & Jones 1999). In the auditory domain, humans show improvements in pitch comparison (Jones et al., 2002), target sound detection (Rimmele et al., 2011), and reaction time (Rimmele et al., 2011) when acoustic stimuli are isochronous vs. aperiodic. In complex acoustic environments, like the ocean, rhythm in vocalizations could enhance perception and processing of acoustic signals and direct attention in an energetically economical way. In humans, joint attention can facilitate synchronization of brain activity across individuals in a group (Dikker et al., 2017), which can in turn enable behavioral synchronization(Sebanz et al., 2006).

Interpersonal synchrony in humans has a litany of prosocial consequences, including increasing trust (Anshel & Kipper, 1988), cooperation (Wiltermuth & Heath, 2009), interpersonal affiliation (Hove & Risen, 2009), helping behavior (Kirschner & Tomasello, 2010), social bonding (Tarr et al., 2015), closeness (Wiltermuth & Heath, 2009), and even pain thresholds during group physical activities (Tarr et al. 2015). This plethora of benefits suggests a strong evolutionary push to develop individual rhythmic faculties that enabled synchronization throughout human history (McNeil, 1997).

Synchronous displays are not limited to humans, but are also used by a diversity of taxa to attract mates, confuse or deter predators, advertise territory, and signal coalitions (Ravignani, 2019a). Theoretical and empirical work show that rhythm in general, and isochrony in particular, is a powerful tool for achieving synchronization because it makes the timing of upcoming signals predictable (Di Paolo, 2000; Merker et al., 2009; Bowling et al., 2013). Behavioral synchronization in cetaceans is common and occurs in various contexts and over various time scales (e.g. diving, Aguilar de Soto et al., 2020; breathing, Sakai et al., 2010; swimming Senigaglia et al., 2012). In a related vein, many cetacean species are highly social and have complex societies; vocal rhythm may have been evolutionarily selected for in cetaceans precisely because, as in humans, it enables synchronization and subsequent prosocial benefits (Moore et al., 2020).

Collectively, these examples demonstrate how vocal rhythm can influence sociality by modulating memory, attention, and synchronization. Given that many cetacean species learn and remember diverse vocal repertoires and complex vocal displays; inhabit a dynamic acoustic environment; and use synchronization to affirm social ties and in various behavioral contexts, the potential adaptive advantages for vocal rhythm are clear, and may explain why it is prevalent across cetacean species.

### *2.5.2. – Isochrony and heterochrony*

Isochronous rhythms abound in the acoustic signals of many species, including bonobos (*Pan paniscus*; De Waal, 1988), zebra finches (Norton & Scharff, 2016), harbor seals (*Phoca vitulina*; Ravignani et al., 2019), and humans (Fitch, 2012). Accordingly, in both mysticete and odontocete vocalizations, isochrony appears to be more common than heterochrony.

From a signal processing perspective, isochronous signals are powerful because they are deterministic, predictable, and economical (Merker et al., 2009; Ravignani & Madison, 2017). Through isochronous repetition, a signaler minimizes the entropy and maximizes the redundancy of a signal (Handel et al., 2012; Ravignani & Madison, 2017). As noted in Kershenbaum et al. (2016), environmental conditions can drive trade-offs between temporal and spectral resolution in vocalizations and impact rhythm development. For many marine species, including cetaceans, sound is the primary modality for communication, but this communication takes place in an acoustically complex environment. Temporal features of vocalizations are more robust to propagation effects and transmission loss than spectral features. For example, the rhythmic features of vocalizations remain constant or nearly constant even as an animal's orientation or distance to a sound receiver—be it a conspecific or hydrophone—changes. This phenomenon is exploited in passive acoustic analyses and likely by the animals themselves (André & Kamminga, 2000; Le Bot et al., 2013; Zaugg et al., 2013). In contrast, the frequency distribution of vocalizations can be highly modulated by characteristics of the transmission environment and the animal's behavior. Encoding information in the rhythmic aspects of vocalizations could thus be an anti-masking strategy, preserving transmission fidelity despite background noise/interference or vocalizer movements (Brumm & Slater, 2006; Zaugg et al., 2013; Schneider & Mercado III, 2019).

Isochronous signals also generate temporal expectancies about upcoming signal events, which can facilitate synchronization among signal producers and/or receivers. Isochrony may itself be a by-product of the fundamental need to synchronize—a need that is apparent across taxa—which could explain the prevalence of isochrony in vocalizations from diverse species (Fitch, 2012; Ravignani et al., 2014). Heterochronous signals can also generate temporal expectancies, but simple rhythms are generally easier to track and synchronize to than complex rhythms, at least in humans (Jones & Pfordresher, 1997; Drake et al., 2000).

The same features that make isochronous signals optimal for transmission fidelity and predictability, however, leave little room for expressivity (i.e. "the capacity of a signaling system to convey different meanings"; Ravignani & Madison, 2017). Expressivity can be achieved by varying the spectral or temporal features of calls, or both. For species where the temporal patterning of signals may be more important than the acoustic units themselves, such as sperm whales, heterochronous signals may have evolved to enable expressivity (Ravignani & Madison, 2017). In the eastern tropical Pacific, sperm whales from the 'Regular', 'Plus-One', and 'Four-Plus' clans do not interact, implying that they have some way of discerning 'us' from 'them' (Rendell & Whitehead, 2003b). There is evidence (see chapter four) that certain coda types with distinct rhythmic patterns act as symbolic markers of cultural identity in sperm whales, as distinct vocal rhythmic patterns can in humans (Jacoby & McDermott, 2017).

### *2.5.3 – Trends in mysticete vocal rhythm*

While singing is well documented in humpback, blue, fin, minke, and bowhead whales (Janik, 2009), research suggests that other mysticetes (many of which are understudied) sing as well (e.g. Omura's whales, Cerchio et al., 2015; North Pacific right whales, Crance et al., 2019; Bryde's whales, Oleson et al., 2003; sei whales, Tremblay et al., 2019). Given that mysticete song is sung exclusively by males in the species for which singer sex is known (e.g. Croll et al., 2002), it is hypothesized to play a role in reproduction, either by mediating male/male competition and/or by attracting females. To understand why rhythm in mysticete vocalizations has only been quantified to date in song, we need to consider (from the perspective of both the song producer and song receiver) why rhythm is a useful feature in song and, contrastingly, why it may be less useful in non-song vocalizations.

Several researchers have hypothesized that rhythm may act as a mnemonic device in humpback whale song, helping males learn and remember the complex and changing vocal displays (Guinee & Payne, 1988; Handel et al., 2009; Green et al., 2011). Guinee and Payne (1988) found that rhythmically predictable themes (i.e. those with static/unchanging phrases) were the most abundant theme type in humpback whale songs and that there was more rhythm in the parts of songs that were most likely to change from year to year. In bowhead whales, complex songs ("diverse sequences of simple and complex calls") are more stereotyped and have more constant repetition patterns than simple songs (simple moan sequences), providing another link between rhythm and song complexity (Delarue, Laurinolli, et al., 2009). Song producers may thus benefit from the memory-enhancing capabilities of rhythm.

Rhythm can also help receivers segment incoming signals (Cutler, 1994) and direct their attention (Gedamke, 2004). Recent research found that male zebra finch song has an underlying isochronous pulse (Norton & Scharff, 2016). Female zebra finches may use periodic attention when listening to male song, attending to pitch at isochronous time points to assess male consistency (Norton & Scharff, 2016). As in humans, rhythm may thus allow female zebra finches to optimize their attention timing and arousal levels when a male is performing (Huron, 2008; Norton & Scharff, 2016). In an acoustically complex environment like the ocean, rhythmicity in mysticete songs may function in a similar way by guiding receivers' attention to salient song features (Smith, 1991).

Rhythmic signals can influence sociality by providing regular updates on a signaler's location, behavioral state, or population of origin. In the context of mysticete song, this could help females locate a male of interest and/or help males mediate their interactions with each other. Singing dwarf minke whales, whose songs are composed of isochronous or heterochronous 'star wars vocalizations', moved away from a speaker during song playback experiments and increased the tempo of their songs (Gedamke, 2004). This suggests that song tempo, and possibly other song features, serve a spacing function in this species (Gedamke, 2004). Song rhythm and tempo may also provide cues about a singer's motivational state, as seen in the threat calls of a passerine bird, the Arabian babbler (*Argya squamiceps*); during clashes with other groups, aggressive males emit isochronous signals, while more timid males emit quieter, more aperiodic signals (Zahavi & Zahavi, 1999). Lastly, song rhythm and tempo can provide cues regarding a singer's population of origin (e.g. blue whales, McDonald et al., 2006; minke whales, Oswald et al., 2011; fin whales, Širović et al., 2017), which has important implications if preferences exist for mating or interacting with individuals from the same or different populations (Delarue, Todd, et al., 2009).

In the acoustically complex ocean, rhythm may thus play a functional role in mysticete song by helping singers learn and remember songs; allowing receivers to predict and attend to salient song features at precise moments in time; and providing cues to the location, behavioral state, and/or population origin of the singer. With so many potential benefits, why then is rhythm less abundant in non-song vocalizations? For example, while dwarf minke whale song has rhythmic structure, social sounds do not (Gedamke, 2004). Similarly, blue whales produce A-B call pairs (comprised of a pulsed A call followed by a tonal B call) both within and outside of song, but the interval between A-B call pairs in song is much more isochronous than in non-song vocalizations (Oleson et al., 2007). There are several possible explanations for this discrepancy, including vocalization complexity, vocalization function/scale, and research effort.

For several mysticete species, like humpback and bowhead whales, songs are the most complex vocalization in the vocal repertoire, and rhythm may be essential in helping singers learn and remember these complex displays. Rhythm could thus be more prevalent in song compared to non-song vocalizations because song is more complex and challenging to produce/remember. However, other mysticete species' songs can be quite simple (in terms of both units and silent intervals, e.g. fin whale song, Širović et al., 2017), which implies that signal complexity alone cannot explain why non-song rhythm is less abundant than song rhythm. Evidence suggests that mysticete song is primarily a long-range communication signal (Mooney et al., 2016; Širović et al., 2007), whereas social sounds are more often made in gregarious aggregations (Gedamke, 2004; Rekdahl et al., 2015). Close proximity to other animals may render some of the aforementioned benefits of vocal rhythm, like localization, less beneficial or necessary. Finally, song has been studied more extensively than non-song vocalizations in many mysticetes (although this is changing, e.g. Rekdahl et al., 2015) and much of the research on mysticete vocalizations has focused on the spectral/combinatorial domain rather than the temporal/rhythmic domain (Clark, 1998). As research on non-song vocalizations (and specifically their temporal features) increases, examples of rhythm in these vocalizations may increase as well.

### *2.5.4 – Trends in odontocete vocal rhythm*

In contrast to mysticetes, rhythm is prevalent across vocalization types in odontocetes. This prevalence likely relates to the evolution of echolocation. During echolocation, animals produce clicks or whistles and use the returning echoes to sense and track features of their environment, including prey. In odontocetes, echolocation evolved approximately 28 million years ago (Geisler et al., 2014) to exploit untapped food niches with active prey. With the evolution of echolocation came selection for an extraordinary auditory system.

The temporal resolution of the odontocete auditory system rivals that of most mammals (Mooney et al., 2009). Studies on captive and wild animals have provided ample evidence that odontocete brains respond isochronously to isochronous stimuli and can follow very rapid series of clicks (e.g. Mooney et al., 2009; Popov & Supin, 1998; Szymanski et al., 1998). Given that sound travels approximately five times faster in water than in air, marine echolocators need an auditory system with high temporal resolution to rapidly interpret returning echoes in an acoustically cluttered environment. This process may lend itself to isochrony because there is an upper limit on how quickly clicks can be produced while still being able to interpret the returning echo (i.e. the two-way sound transit time; Backus & Schevill, 1966; Ridgway, 2011). A similar limit is seen in the greater sac-winged bat (Burchardt et al., 2019). During foraging, echolocating isochronously at or near the two-way sound transit time could allow odontocetes to receive the most regular updates on prey location possible without self-masking from returning echoes. Similar anti-masking strategies may underpin the call timing observed in some conspecific vocal exchanges (e.g. bottlenose dolphin signature whistle exchanges, Nakahara & Miyazaki, 2011; sperm whale coda exchanges, Schulz et al., 2008).

There is neurological evidence reinforcing the link between attention and rhythm in odontocetes. Bottlenose dolphins dampen their hearing sensitivity when acoustically warned about incoming loud sounds (Nachtigall & Supin, 2015). This dampening is *lessened* when the loud sounds occur predictably after the warning sounds compared to unpredictably (Nachtigall et al., 2016). Predictability, which can be achieved through isochrony, thus lets bottlenose dolphins precisely pinpoint when their hearing should be dampened, allowing them to maintain a full auditory scene for longer (Nachtigall et al., 2016). As in other species, rhythm could therefore allow odontocetes to modulate attention in energetically economical ways.

The anatomical and neurological features that enable echolocation likely underpin the prevalence of rhythm in other types of odontocete vocalizations. Four of the most extensively studied odontocete species—common bottlenose dolphins, Indo-Pacific bottlenose dolphins, orcas, and sperm whales—show isochronous rhythm in vocalizations used in different behavioral contexts, suggesting that, at least for these species, rhythm is a fundamental feature of communication. Similarly, the greater sac-winged bat—which, like many odontocetes (Tyack & Sayigh, 1997), is a vocal learner—exhibits isochronous rhythm in various call types (Burchardt et al., 2019). The rhythmic faculties gained with the evolution of echolocation may have been exapted to serve additional functions, including conveying identity (be it species, population, cultural, or individual) and potentially facilitating or reinforcing behavioral synchronization (Herzing, 2015). Given the current trends, targeted research on the rhythmic properties of vocalizations made by less well-studied odontocetes will likely unearth more examples of rhythm in additional behavioral contexts.

### *2.5.5 – Why do mysticetes and odontocetes differ?*

That mysticetes and odontocetes appear to differ in the prevalence of rhythm across different vocalization types is interesting and warrants further consideration. Anatomy, foraging strategy, and social structure could all play a role in the disparity.

The sound production mechanisms in mysticetes and odontocetes are fundamentally different. Mysticetes vocalize by vibrating the folds of the larynx, while odontocetes vocalize by forcing air through the phonic lips (Reidenberg & Laitman, 2018). These anatomical variations relate to foraging strategy, with odontocetes using echolocation to target single active prey while mysticetes use other senses to find schooling prey. As previously mentioned, the evolution of echolocation and corresponding high-resolution auditory system is likely a driving factor behind the prevalence of rhythm across vocalization types in odontocetes. With mysticete foraging vocalizations generally rare (and, when present, used to coordinate behavior or startle rather than detect prey; Baker, 1985; Sharpe, 2001), it is intuitive that *rhythmic* foraging vocalizations are rare as well (but see humpback whale foraging cries for a potential counterexample; Cerchio & Dahlheim, 2001). This scarcity could, however, also reflect a paucity of research on and knowledge of foraging vocalizations in mysticetes.

Relevant to rhythm, trends in the hierarchical temporal structure of mysticete and odontocete vocalizations suggest a role for social structure. Kello et al. (2017) found that hierarchical temporal structure in vocalizations is enhanced by social interactions across species. Orca vocalizations, which are used in vocal interactions, had more hierarchical temporal structure (on par with that of human conversations) than humpback whale song, which is sung by solitary males (Kello et al., 2017). The abundance of rhythm in odontocete vocalizations may therefore reflect their generally more complex societies compared to mysticetes, highlighting known links between vocal and social complexity (Freeberg et al., 2012; McComb & Semple, 2005; Ord & Garcia-Porta, 2012).

### *2.5.6 – Limitations and recommendations*

Across cetacean species, and for mysticetes in particular, significantly more research has focused on spectral features of vocalizations than on temporal features (reflecting “a history of infatuation with melodic qualities”; Clark, 1998). The study of cetacean vocal rhythm thus remains ripe with research avenues and opportunities for expansion in the future but would benefit from researchers being explicit in the language they use to describe rhythmic phenomena.

Inconsistent terminology constituted a significant obstacle in understanding, interpreting, and synthesizing the current state of knowledge on cetacean vocal rhythm. These inconsistencies also made a fully systematic review of the literature infeasible because publications that assessed temporal features of cetacean vocalizations often did so without invoking the word rhythm (Method S2.1 in Appendix A; Figures S2.1/S2.2). Additionally, terms like tempo and rhythm are often conflated in cetacean literature (e.g. André & Kamminga, 2000) and beyond (Fitch, 2013) but are distinct features of acoustic signals (McAuley, 2010) and should be referenced as such. Definitions of rhythm are also sometimes restricted to isochrony (e.g. Herzing, 2015), which inherently limits what qualifies as rhythm. We recommend that researchers converge on a definition of rhythm that allows for different types of rhythm and is sufficiently broad to be applicable across species, like the definition we have used in this paper.

Our definition of rhythm is not restricted to an explicit time scale, so future studies should clearly describe the time scale analyzed (as perceived signal rhythmicity can vary depending on the temporal window) and consider looking for rhythmic patterns across multiple scales (e.g. Kello et al., 2017). Defining such scales in terms of number of calls (e.g. 100 consecutive calls) rather than specific periods of time (e.g. 100 minutes) could facilitate more informative comparisons, for example among species with order of magnitude differences in inter-call intervals (e.g., fin whale song vs. Omura's whale song; Cerchio et al., 2015; Širović et al., 2017). Furthermore, we recommend that researchers center their descriptions of vocal rhythm within frameworks like that from Ravignani et al. (2014) to make descriptions more consistent, interpretable, and comparable across studies.

How researchers designate the acoustic unit of interest also has strong implications for if and what type of rhythm is detected, and best practices for making that designation are not always clear (Sainburg et al., 2020). For example, a bout of sperm whale '5R' codas (i.e. codas with five isochronously spaced clicks) exhibits isochrony if the unit of interest is each coda (with equal spacing between each coda), but heterochrony if the unit of interest is each click (because the inter-click intervals within the coda are shorter than the inter-click intervals separating two codas) (Figure 2.2).

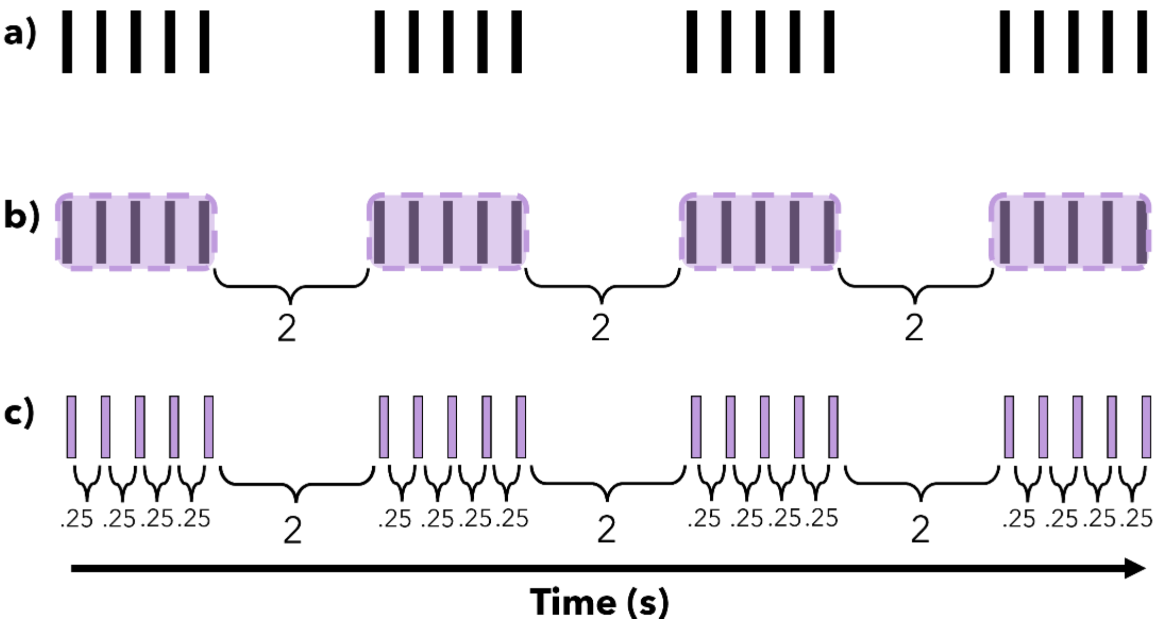


*Figure 2.2* – Schematic illustrating how detected rhythm can change based on acoustic unit of interest designation. Each vertical bar represents one sperm whale coda click and the purple shading denotes the acoustic unit of interest. Black brackets denote silent intervals, with toy durations given below. (a) A bout of four '5R' codas can be divided into acoustic units in two ways: (b) if each coda is considered the unit of interest, an isochronous pattern emerges (with equal spacing of 2 s between each coda in the bout); (c) if each click is considered the unit of interest, a heterochronous pattern emerges (with four short intervals of 0.25 s followed by a longer interval of 2 s).

Ideally, our designation of the acoustic unit of interest would be based on the animals' perception, but detailed perceptual information is often lacking for cetaceans. One way to approach this challenge is by considering different units of interest within the same vocalization and quantifying rhythm for each unit, as we and others have done (e.g. Herzing, 2015). An alternative is to consider any sound flanked by silences as a unit of interest (e.g. Burchardt et al., 2019), but this approach can be limiting for hierarchically organized vocal displays like humpback whale song because it negates consideration of higher-order rhythm (e.g. among phrases or themes). The method used by Kello et al. (2017) to measure hierarchical temporal structure in vocalizations across different timescales could help deal with this issue. Regardless of the approach taken, future vocal rhythm research should include explicit discussions of the acoustic unit of interest.

Technically speaking, much of the evidence of cetacean vocal rhythm presented in this review was inferred from relatively simple and straightforward metrics. At the very least, cetacean researchers should provide the CV of the inter-unit intervals in future work, as this allows a first assessment of whether a vocalization might be isochronous. An abundance of techniques exist for quantifying vocal rhythm (reviewed in Ravignani & Norton, 2017) and have already been illuminating in studies of cetaceans (e.g. Schneider & Mercado III, 2019). Researchers will be able to describe vocal rhythm in more nuanced ways if they expand their toolboxes to include some of these techniques.

### *2.5.7 – Future research directions*

Many papers we encountered provided preliminary evidence of rhythm in mysticete (e.g. blue whale arch sound sequences, Mellinger & Clark, 2003; sei whale song, Tremblay et al., 2019) and odontocete (e.g. orca disconnected multi-loop whistles, Riesch & Deecke, 2011; short-finned pilot whale repeated call sequences, Sayigh et al., 2013) vocalizations (Tables S2.1/S2.2). This abundance of preliminary evidence suggests that there are many more examples of cetacean vocal rhythm production than we have documented here, and we encourage targeted rhythm research on these call types.

To understand the adaptive value of vocal rhythm, a logical next step after quantifying rhythm in vocalizations from single animals is to quantify rhythm in duets and/or choruses using techniques like cross-correlation, multidimensional time series, or circular statistics (Ravignani & Norton, 2017). Such research will help clarify the links between vocal rhythm and behavioral synchronization in cetaceans, as it has in other species (e.g. plain-tailed wrens, *Thryothorus euophrys*; Mann et al., 2006). Once again, the definitional framework from Ravignani et al. (2014) (i.e. the “chorus tree”) can help guide and standardize investigations.

Another interesting area of research relates to the ability to perceive and synchronize body movements (such as clapping, wingbeats, head bobs, or fluke beats) to an external beat (Wilson & Cook, 2016). Humans find it difficult to *not* engage in beat perception and synchronization, but this proclivity has, until recently, seemed largely absent in the animal kingdom (Fitch, 2012; Wilson & Cook, 2016). Using fine-scale sound and movement recording tags, coupling between motor and vocal output can be investigated in groups of cetaceans. Odontocetes are particularly interesting candidates for such studies, given their capacity for behavioral synchronization and the evidence that various bat species couple wingbeats to echolocation and respiration during flight (e.g. Burchardt et al., 2019; Suthers et al., 1972).

Future research could also look more in-depth for trends in heterochronous rhythm in cetacean vocalizations. For odontocetes that primarily communicate using clicks instead of whistles, such as sperm whales, harbor porpoises, and beaked whales, we might expect to see more examples of heterochronous rhythm because expressivity may be less likely to be achieved by modulating spectral parameters of calls. Heterochronous rhythm may also be more common in calls that relate to or convey identity compared to, for example, foraging calls, because multiple identities require multiple distinct signals (Bouchet et al., 2013). Rhythm in sperm whale echolocation vs. codas provides very preliminary support for this hypothesis (Table 2.3). Further exploration of rhythmic trends in calls that serve different functions will help us move beyond merely documenting rhythm and towards a more holistic understanding of how rhythm is being utilized.

## 2.6 – Conclusions

Vocal rhythm can augment sociality and behavior in nontrivial ways but, until recently, has been a taxonomically restricted field of research. By increasing the diversity of species represented in the field and embracing a comparative approach, we can better understand the evolutionary history and adaptive value of vocal rhythm. In summarizing what is currently known about rhythm in cetacean vocalizations, this work represents a first grab at one of the many “low-hanging fruits” in comparative rhythm research (Ravignani, Dalla Bella, et al., 2019). By quantifying rhythm at various scales and in diverse species, expanding our technical toolbox, and being more intentional in how we describe vocal rhythm, cetacean researchers specifically, and bioacousticians more broadly, can contribute to a rapidly growing and cross-species body of work on vocal rhythm.

## 2.7 – Acknowledgements

This work greatly benefited from discussions with Noor Youssef, Andrea Ravignani, Alexander South, Andrew Horn, Shane Gero, Hilary Moors-Murphy, Aaron

Newman, Luke Rendell, Connor Reid, and Koen de Reus. Additional thanks to Luke Rendell for hosting TAH at the University of St. Andrews. We are grateful to the researchers and organizations who answered questions about their work, sent supplemental data, and directed us to additional resources. Specifically, we thank: Anne Simonis and the Scripps Whale Acoustic Lab (University of California, San Diego), Christopher Tremblay, Ellen Garland, Elizabeth Zwamborn, Emily Doolittle, Erin Oleson, Jason Gedamke, Jessica Crane, Julien Delarue, Kate Stafford, Katie Kowarski, Laura Feyrer, Patrick Miller, Rebecca Wellard, Rudiger Riesch, Sam Walmsley, Shannon Rankin, Sofie Van Parijs, and Susan Parks.

# CHAPTER 3

# USING IDENTITY CALLS TO DETECT STRUCTURE IN ACOUSTIC DATASETS[16]

> "Wildness has its own understandings."
>
> —J. B. Mackinnon

## 3.1 – Abstract

Acoustic analyses can be powerful tools for illuminating structure within and between populations, especially for cryptic or difficult to access taxa. Acoustic repertoires are often compared using aggregate similarity measures across all calls of a particular type, but specific group identity calls may more clearly delineate structure in some taxa. We present a new method—the identity call method—that estimates the number of acoustically distinct subdivisions in a set of repertoires and identifies call types that characterize those subdivisions. The method uses contaminated mixture models to identify call types, assigning each call a probability of belonging to each type. Repertoires are hierarchically clustered based on similarities in call type usage, producing a dendrogram with 'identity clades' of repertoires and the 'identity calls' that best characterize each clade. We validated this approach using acoustic data from sperm whales, grey-breasted wood-wrens, and Australian field crickets, and ran a suite of tests

---

[16] This chapter has been published in the journal *Methods in Ecology and Evolution* as: Hersh, T. A., Gero, S., Rendell, L., Whitehead, H. (2021). Using identity calls to detect structure in acoustic datasets. *Methods in Ecology and Evolution*, 12(9), 1668–1678. doi: 10.1111/2041-210X.13644. TAH and HW conceived and designed the method, with input from SG and LR. TAH analyzed and interpreted the data and drafted the manuscript. All authors collected sperm whale coda data, revised the manuscript, and read and approved the final version. The manuscript was received by the journal on September 7, 2020; resubmitted on March 17, 2021; accepted on May 5, 2021; published online (standalone) on May 20, 2021; and published online (with issue/volume) on September 1, 2021. *Methods in Ecology and Evolution* is owned by John Wiley & Sons Ltd. As per their website (https://www.wiley.com/network/researchers/latest-content/how-to-clear-permissions-for-a-thesis-or-dissertation, accessed on December 8, 2021), "If you are the author of a published Wiley article, you have the right to reuse the full text of your published article as part of your thesis or dissertation. In this situation, you do not need to request permission from Wiley for this use." See Copyright Release S3.1 in Appendix E for additional details.

to assess parameter sensitivity. For all taxa, the method detected diagnostic signals (identity calls) and structure (identity clades; sperm whale subpopulations, wren subspecies, and cricket species) that were consistent with past research. Some datasets were more sensitive to parameter variation than others, which may reflect real uncertainty or biological variability in the taxa examined. We recommend that users perform comparative analyses of different parameter combinations to determine which portions of the dendrogram warrant careful vs. confident interpretation. The presence of group-characteristic identity calls does not necessarily mean animals perceive them as such. Fine scale experiments like playbacks are a key next step to understand call perception and function. This method can help inform such studies by identifying calls that may be salient to animals and are good candidates for investigation or playback stimuli. For cryptic or difficult to access taxa with group-specific calls, the identity call method can aid managers in quantifying behavioral diversity and/or identifying putative structure within and between populations, given that acoustic data can be inexpensive and minimally invasive to collect.

## 3.2 – Introduction

The natural world is inherently noisy, and animals that communicate acoustically must evolve ways of ensuring signal transmission fidelity in the face of such noise. Variation in acoustic signals can provide information about the identity of individuals, groups, populations, and species, as well as drivers of signal evolution (Wilkins et al., 2013). Biologists who want to understand animal communication systems need robust, efficient ways of classifying and comparing acoustic signals.

Animal calls (i.e. acoustic signals) and repertoires (i.e. collections of calls) have often been compared using measures of similarity calculated over a range of call parameters. However, the presence of group-specific calls in some taxa (e.g. canids, Kershenbaum et al., 2016; sperm whales, Rendell & Whitehead, 2003) may indicate various levels of biological structure (e.g. social groups, cultures, subpopulations, populations, subspecies, species, etc.) to human observers more clearly than comparisons of aggregate similarity measures across calls or repertoires. This is especially likely if those group-specific calls—which we term ‘identity calls’—have been produced by selection pressures for signaling identity, since signal detection theory predicts that calls with such functions will exhibit transmission-enhancing features, including redundancy, stereotypy, and distinctiveness (Wiley, 2013).

We present a new method—the identity call method (hereafter IDcall)—that finds redundant, stereotyped, and distinct identity calls in acoustic datasets and uses those identity calls to predict corresponding biological structure. Calls are first indexed in multivariate space using appropriate quantitative measures. Then, call types are identified by applying parsimonious mixtures of multivariate contaminated normal distributions (hereafter contaminated mixture models) to this multivariate dataset, with each call assigned a probability of belonging to each type. Lastly, repertoires (from single or multiple individuals, Method S3.2 in Appendix B) are hierarchically clustered based on similarities in call type usage. The final dendrogram shows ‘identity clades’ (i.e. clades of repertoires distinguished by certain call types) and the identity calls that characterize each clade. We use ‘clade’ in a non-evolutionary sense to mean a node and all branches descended from it on a dendrogram.

We illustrate the efficacy of this method in finding identity calls that denote previously described biological structure using acoustic data from sperm whales (*Physeter macrocephalus*), grey-breasted wood-wrens (*Henicorhina leucophrys*; hereafter wrens), and Australian field crickets (*Teleogryllus spp.*; hereafter crickets). These datasets were selected to demonstrate the breadth of the method’s capacity to detect different levels of biological structure —sperm whale subpopulations, wren subspecies, and cricket species—using identity calls. In all cases, we knew biological structure existed *a priori* based on past research that often included extensive genetic and/or mark-recapture analyses. Such approaches may not be appropriate or possible for taxa that are cryptic or difficult to access, and are often more costly and logistically challenging than acoustic data collection (Garland et al., 2015). We hope that IDcall can complement such approaches by allowing users to rapidly detect diagnostic signals (identity calls) and putative biological structure (identity clades) that may merit additional research to determine genetic, behavioral, and/or ecological distinctiveness of the animals producing the repertoires, ultimately facilitating more targeted management and conservation action, if necessary.

## 3.3 – Methods

### *3.3.1 – Classifying calls into types*

IDcall was implemented in R version 3.6.1 (R Core Team, 2013). When applicable, the calls were first divided based on the number of quantitative measures (*N*). For example, sperm whale codas (i.e. stereotyped click patterns made in social situations) were divided based on the number of inter-click intervals. All calls, represented as points in *N* dimensional space, were classified into types using contaminated mixture models via the 'ContaminatedMixt' R package (Punzo et al., 2018). For each group of *N*-dimensional calls, we fitted between 2–10 mixture components to the data using the expectation conditional-maximization (ECM) algorithm initialized with the *k*-means algorithm (Punzo & McNicholas, 2016). We fit a range of mixture components to avoid *a priori* specification of the number of call types defined by the resulting components and used *k*-means based on the results of previous simulation studies (e.g. Shireman et al., 2017). To reduce the risk of outliers (i.e. calls that have ambiguous type) being assigned to their own mixture component(s) when fitting higher numbers of mixture components (e.g. Evans et al., 2015), we required all of the fitted models to have outliers by setting the 'contamination' boolean to TRUE in the 'CNmixt' function (Punzo & McNicholas, 2016). The most parsimonious model and number of mixture components was selected based on the Bayesian information criterion (BIC; Schwarz, 1978) for the large sperm whale datasets or the bias-corrected Akaike information criterion (AICc; Hurvich & Tsai, 1989) for the smaller wren and cricket datasets (see section 3.3.3 for dataset descriptions). For each *N*, the final number of mixture components was the number of call types and the call types were validated by comparison to previous research (Method S3).

The contaminated mixture model algorithm estimates the probability that each call, *i*, belongs to each call type, *j*, as $u(i,j)$ (where $u(i,j)=0$ if call *i* has a different *N* from the calls in *j*). We calculated usage, *U*, of each call type, *j*, for each repertoire, *r*, by summing the probability of call type membership for all calls in the repertoire and dividing by the total number of calls in the repertoire, *n*(*r*):

$$U(r,j) = \frac{\sum_{i \in r} u(i,j)}{n(r)}$$

A *U* close to 0 means a call type is rarely used in the repertoire, while a *U* close to 1 means a call type is frequently used. If *U* equals 1, the call type is the only one used in the repertoire.

### *3.3.2 – Delineating identity calls and identity clades*

Average linkage hierarchical clustering was used to cluster repertoires into clades based on quantitative similarity measures between pairs of repertoires, where similarity is calculated as the correlation of call type usages between each pair of repertoires (e.g. correlations of $U(r_1,:)$ with $U(r_2,:)$). Two repertoires that used the same call types with the same relative frequencies would have a high correlation value (close to 1) and cluster close together in the dendrogram, whereas two repertoires that used different call types and/or had opposite trends in call type usage would have a low correlation value (close to -1) and not cluster together. A cluster of repertoires could only be designated as an identity clade if those repertoires were united by high usage of at least one identity call—a call type used frequently by that clade and rarely by all others.

For a call type *j* to be considered an identity call for clade *c*, the mean call type usage across repertoires (mean $U(r \in c,j)$) in a clade had to be greater by a factor of at least a given value (a parameter we call *critfact*) than both the mean call type usage in each other clade with at least a given number of repertoires (a parameter we call *minrep*) at that level of the dendrogram and the mean call type usage of all repertoires not in that clade. Additionally, within a potential identity clade, a call type could only be designated as an identity call if the mean call type usage of the *minrep* repertoires in the clade that used the call type the most was less than *critfact* multiplied by the median call type usage across all repertoires in the clade. This prevented call types used frequently by just a few repertoires in a clade from being designated as identity calls. Extremely rare call types—for which the median usage in the *minrep* repertoires that used the call type the most was 0—were not used to build the dendrogram.

Starting at the base of the dendrogram (where each repertoire is a leaf), we ran an algorithm (Table 3.1) that tested each node in turn to see if it met the criteria to be considered an identity clade (i.e. a clade of at least *minrep* repertoires with at least one identity call in the final tree). In each identity clade, $c$, we calculated the proportion of repertoire calls that were identity calls (the clade identity call proportion, $icp$) from:

$$icp(c) = \frac{\sum_{r \in clade(c)} \sum_{j:\, identity\ call\ for\ clade(c)} U(r,j)}{\sum_{r \in clade(c)} \sum_{j} U(r,j)}$$

To quantify the support for each repertoire belonging to an identity clade, we calculated the correlation between the call type usages of the repertoire and the median usages of the identity clade.

Proximity of identity clades in the dendrograms reflects repertoire similarity but may or may not reflect phylogeny. Two neighboring identity clades generally have repertoires that are more similar to each other than to other clades further away, but that does not necessarily mean they are descended from the same 'ancestor repertoire'. An understanding of how the calls comprising repertoires change over time is necessary before phylogenetic inferences can be attempted.

*Table 3.1* – IDcall hierarchical clustering algorithm for daughter and parental clades. Each clade has one of three possible statuses: potential identity clade (**P**), final identity clade (**F**), or no classification (**N**; no additional classification occurs). Initially, each repertoire (considered a daughter clade) has status **P**. Clades with status **P** can be **P-** (no identity calls) or **P+** (with identity calls; **x** and **y** denote different identity call types). Statuses of daughter and parental clades (the clades being merged and the resultant clade, respectively) change at each node following the rules below. When all clades have status **F** or **N**, or the correlation between two clades joined at a node is less than zero, the process terminates.

| **Daughter clade 1 status** | **Daughter clade 2 status** | **Result** |
|---|---|---|
| P- | P+ | Parental clade becomes P+ (has identity calls) or P- (does not have identity calls) |
| P- | P- | |
| P+$_x$ | P+$_y$ | Both daughter clades become F and are not merged; parental clade is N |
| F | P+ | Daughter clade that was P+ becomes F; parental clade is N |
| N | P+ | |
| F | P- | Daughter clade that was P- becomes N; parental clade is N |
| N | P- | |
| F | F | Parental clade is N |
| N | N | |
| F | N | |

### *3.3.3 – Test datasets*

We used acoustic data from sperm whales, wrens, and crickets to test the ability of IDcall to find previously described biological structure. Our analyses are illustrative but should not be regarded as definitive results for each taxon.

Globally, sperm whale populations are divided by the geography of the different oceans, where only males move genes between basins (Lyrholm et al., 1999), but can also be delineated locally into sympatric cultural clans which are defined based on the codas they use (Rendell & Whitehead, 2003). Clans represent socially segregated subpopulations that are phenotypically distinct (Rendell & Whitehead, 2003). The sperm whale dataset contains 13,805 codas recorded in the Atlantic/Mediterranean (Table S3.1, Figure S3.1) and 18,481 codas recorded in the Pacific (Table S3.2, Figure S3.1). Our objective was to determine if different clans could be distinguished by 'identity codas' using IDcall applied to coda absolute inter-click intervals ($N$=2–8). Codas were divided into repertoires based on the group identity of the recorded whales (Method S3.2 in Appendix B).

The wren dataset (Halfwerk et al., 2016; Dryad Digital Repository https://doi.org/10.5061/dryad.q5p7g) contains 396 averaged song types from males of two subspecies (*H. l. hilaris* and *H. l. leucophrys*) whose ranges overlap but eventually replace each other with changing altitude in the Ecuadorian Andes. These subspecies are genetically distinct but morphologically similar and can be delineated based on their characteristic song (Dingle et al., 2008). Our goal was to determine if IDcall could distinguish 'identity songs' of the two subspecies using several song frequency parameters (averaged note peak frequency, minimum and maximum song frequency; $N$=3). Songs were divided into repertoires by individual (n=41 males; Method S3.2 in Appendix B).

The cricket dataset (Moran et al., 2020; Dryad Digital Repository https://doi.org/10.5061/dryad.wpzgmsbhr) is comprised of male calling song data for two closely-related species (*Teleogryllus commodus,* 127 songs; *Teleogryllus oceanicus,* 131 songs) which live in sympatry across hundreds of kilometers of the Australian east coast and show similar habitat and resource use (Moran et al., 2020). The two species produce hybrids in the lab but rarely or not at all in the wild, with acoustics helping prevent wild crossbreeding (Moran et al., 2020). Our objective was to see if IDcall could distinguish *T. commodus* and *T. oceanicus* ‘identity songs’ using several interval-based song features (chirp pulse length, chirp interpulse interval, chirp-trill interval, and trill-pulse length; *N*=4). Songs were divided into repertoires by field site (n=16 sites; Method S3.2 in Appendix B).

### *3.3.4 – Testing options/parameters and comparing dendrograms*

Several options and parameters must be set in IDcall. To assess how varying each impacted the final identity calls and clades, we tested a range of settings/values, changing one at a time while keeping the others at established defaults (Table 3.2), and compared each ‘trial dendrogram’ to a ‘baseline dendrogram’. The default values should not be interpreted as optimized values for each dataset but provided a reasonable starting point to assess variation across trials; other settings may be more appropriate for other datasets (see section 3.5.5 for recommendations). To assess start-point dependence during call classification, we reran the default parameters twice and compared both dendrograms to the baseline dendrogram.

Within and across datasets, and for each trial (n=19 per dataset), we assessed how the number of call types (total and identity), number of identity clades, and the tree identity call proportion varied. We also compared the repertoire composition of identity clades in each trial to the baseline by looking at all possible repertoire pairs and determining if each pair was assigned to the same or different clade in the two dendrograms. All repertoires not assigned to an identity clade were considered part of a single ‘outlier clade’. Similarity was calculated as the number of repertoire pairs assigned to the same clade in both dendrograms plus the number of pairs assigned to different clades in both dendrograms divided by the total number of pairs. To verify that the method was not prone to detecting spurious identity calls and clades, we randomly assigned calls to repertoires ten times for each dataset and used the default values to see if identity calls and clades were still delineated.

*Table 3.2* – Default and trial values for IDcall options and parameters. Default settings were used across datasets unless noted. In the wren and cricket datasets, slightly smaller *minrep* values were tested and AICc was used as the default information criterion to account for the smaller dataset sizes.

| **Stage** | **Option/ Parameter** | **Default setting** | **Additional tested settings/values** | **General method behavior** |
|---|---|---|---|---|
| Call classification | ECM initialization strategy | *k*-means | random.post, random.clas | Effect on number of call types (total and identity) and tree identity call proportion varied across datasets but typically yielded similar trees. |
| | Information criterion | BIC (sperm whales)<br>AICc (crickets, wrens) | AIC, ICL, AICc (sperm whales)<br>AIC, ICL, BIC (crickets, wrens) | Using AIC/AICc produced the most call types. Using ICL produced the least, with BIC intermediate. |
| Hierarchical clustering | Linkage method | average | single, complete | Single linkage resulted in chained trees with long, straggly clusters. Complete linkage resulted in trees with compact clusters but more repertoires designated as outliers. Average linkage trees were more similar to single linkage trees than to complete linkage trees. |
| | *critfact* | 14 | 6, 10, 18, 22, 26 | Increasing *critfact* corresponded with a decreasing number of identity calls/clades and tree identity call proportion. |
| | *minrep* | 6 | 4, 8, 10, 12, 14 (sperm whales)<br>3, 5, 7, 9, 11 (crickets, wrens) | Increasing *minrep* corresponded with a decreasing number of identity calls/clades. Effect on tree identity call proportion varied across datasets. |

## 3.4 – Results

### *3.4.1 – General method performance*

For all four datasets, IDcall found identity calls and delineated identity clades (for baseline dendrogram results, see Table S3.3). The Atlantic/Mediterranean sperm whale, wren, and cricket dendrograms were robust (in terms of number and composition of identity clades) across trials, while the Pacific sperm whale dendrograms exhibited more variability (Tables S3.4–S3.7, Data S3.1–S3.4 in Appendix B). When the repertoire grouping variable of each call was randomly permuted, no identity calls or clades were delineated in any dataset.

As *critfact* increased, the number of identity calls and tree identity call proportion generally decreased. The number of identity clades did not change across trials for Atlantic/Mediterranean sperm whales or crickets, but gradually decreased for Pacific sperm whales and decreased at the highest tested value for wrens.

With increasing *minrep*, the number of identity clades generally decreased. For the sperm whale and cricket datasets, the number of identity calls decreased as well. The opposite trend was seen for wrens. The tree identity call proportion decreased for Pacific sperm whales and crickets but increased for Atlantic/Mediterranean sperm whales and wrens. When *minrep* was increased beyond the expected identity clade size, no identity clades were delineated.

Using random.post (i.e. random generation of the initial matrix with posterior probabilities of group membership; Punzo et al., 2018) for ECM initialization generally increased the number of call types (total and identity) compared to the baseline for sperm whales but had the opposite effect for wrens/crickets. With random.clas (i.e. random generation of the initial classification matrix; Punzo et al., 2018), the effect on the total number of call types varied across datasets, but the number of identity calls and tree identity call proportion decreased for the Atlantic/Mediterranean sperm whale, wren, and cricket datasets (the opposite was seen for Pacific sperm whales). Under both strategies, the number and composition of identity clades was like the baseline except for Pacific sperm whales when using random.post (a baseline dendrogram identity clade was split in two) and for wrens when using random.clas (no identity clades delineated).

Across datasets, using AIC during call classification resulted in the most call types, while ICL resulted in the fewest. Using AIC typically decreased the number of identity calls from the baseline criterion (BIC for sperm whales, AICc for wrens/crickets), but the effect on tree identity call proportion varied across datasets and the final number of identity clades did not change in any dataset. In contrast, using ICL decreased the number of identity calls and the tree identity call proportion, with the number of identity clades remaining constant (sperm whales) or decreasing (wrens and crickets). Start-point dependence during call classification was evidenced by variation in the number of call types (total and identity) in duplicate runs of the default parameters across datasets, but the final identity clades were very similar to those in the baseline dendrograms.

Using single linkage during hierarchical cluster analysis resulted in chained trees with long, straggly clusters, whereas complete linkage produced trees with compact clusters but more repertoires designated as outliers (or no tree at all for wrens). Average linkage and single linkage trees were generally more similar to each other than to complete linkage trees.

### *3.4.2 – Sperm whales*

For most parameter combinations (n=16/20, including the baseline), the Atlantic/Mediterranean sperm whale coda repertoires were delineated into three identity clades corresponding to the two known eastern Caribbean clans (EC1 and EC2; Gero et al., 2016) and the known Mediterranean clan (Rendell & Frantzis, 2016) (for baseline dendrogram, see Figures 3.1/S3.2; average trial dendrogram similarity=0.992±0.017). The Gulf of Mexico and Panama repertoires were outliers. The EC2 clan had the most identity calls (n=5), followed by the Mediterranean (n=3) and EC1 (n=2) (Table S3.8). Only the EC1 and Mediterranean clans were detected at higher values of *minrep*. In the complete linkage dendrogram, all EC1 repertoires were designated as outliers.

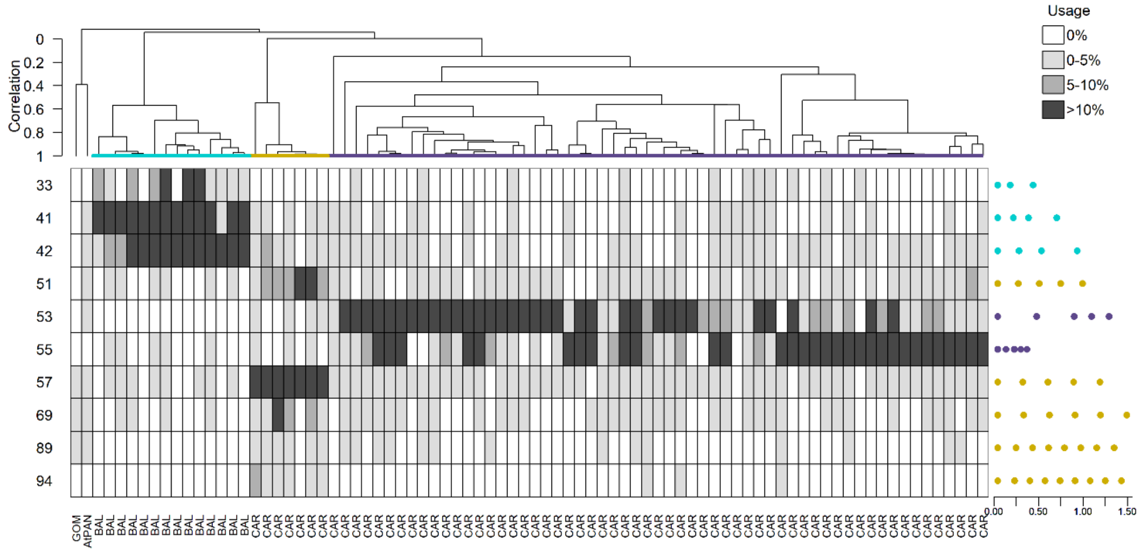


*Figure 3.1* – Baseline dendrogram with clans and identity coda types for Atlantic/Mediterranean sperm whales. Average linkage hierarchical clustering dendrogram (top) depicts similarity among sperm whale coda repertoires recorded in the Atlantic/Mediterranean. Colored identity clades correspond to three known clans: Mediterranean (cyan), EC2 (gold), and EC1 (purple). Heat map (bottom) depicts identity coda type usage (rows) for each repertoire (columns) in shades of grey, which differs from similar figures in past sperm whale studies (e.g. Cantor et al., 2016) because usage here is calculated based on probabilistic assignment of codas to types. Identity coda type codes are on the left (see Table S3.12 for type names) and type centroid rhythm plots (colored by clan) are on the right (each dot represents a click; scale bar is in seconds). Recording locations are listed on the bottom (see Figure S3.1 for abbreviations).

The Pacific sperm whale baseline tree had five identity clades (Figures 3.2/S3.3), four of which correspond to known clans (Regular, Short, Plus-One, and Four-Plus). The putative fifth clan includes repertoires previously considered Four-Plus (Cantor et al., 2016) and a new repertoire. The number of identity codas (2–7) and the clan identity coda proportion (22%–56%) varied across clans (Table S3.9). Results across trials for Pacific sperm whales were more variable, which was reflected in a lower average trial dendrogram similarity (0.972±0.037) than Atlantic/Mediterranean sperm whales. Across trials, the Regular (n=20/20), Plus-One (n=19/20), Four-Plus (n=20/20), and putative fifth (n=16/20) clans were delineated robustly except when *minrep* exceeded the number of repertoires in the clan. The most variability was seen in the baseline Short clan, which often formed one clan (n=13/20) but sometimes two (n=4/20, e.g. Figure S3.4) or three (n=1/20, e.g. Figure S3.5) depending on the trial. The complete linkage dendrogram looked the most distinct from the baseline, with many outlier repertoires (n=34). In the single linkage dendrogram, a smaller version of the Short clan formed, with many of the baseline Short clan repertoires assigned to the Four-Plus clan.

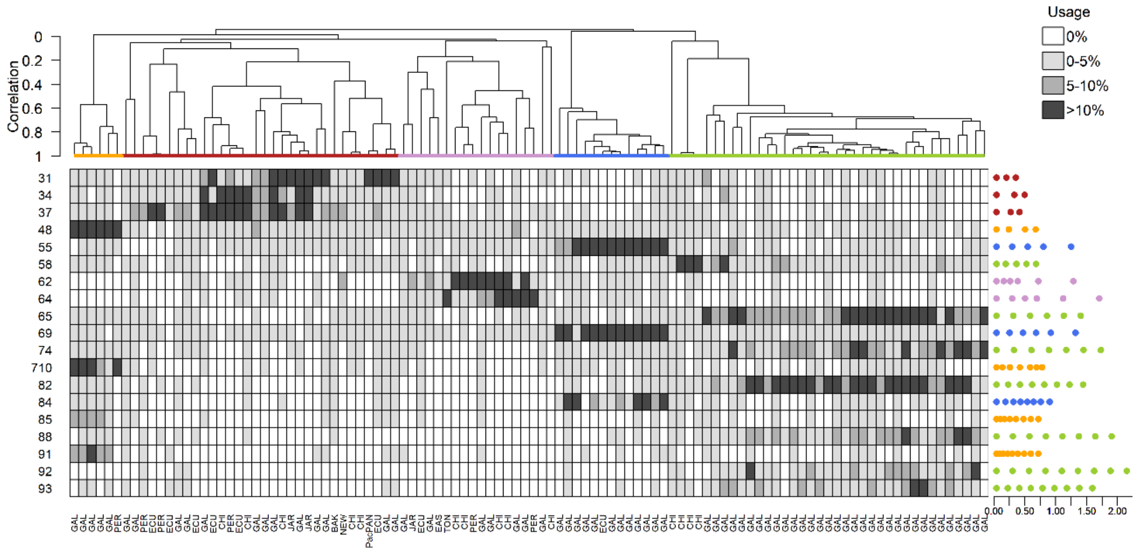


*Figure 3.2* – Baseline dendrogram with clans and identity coda types for Pacific sperm whales. Average linkage hierarchical clustering dendrogram (top) depicts similarity among sperm whale coda repertoires recorded in the Pacific. Colored identity clades correspond to a putative new clan (orange) and four known clans: Short (red), Four-Plus (pink), Plus-One (blue), and Regular (green). Heat map (bottom) depicts probabilistic identity coda type usage (rows) for each repertoire (columns) in shades of grey, which differs from similar figures in past sperm whale studies (e.g. Cantor et al., 2016) because usage here is calculated based on probabilistic assignment of codas to types. Identity coda type codes are on the left (see Table S3.13 for type names) and type centroid rhythm plots (colored by clan) are on the right (each dot represents a click; scale bar is in seconds). Recording locations are listed on the bottom (see Figure S3.1 for abbreviations).

### *3.4.3 – Wrens*

Identity clades were typically delineated in the wren dataset (n=17/20; no identity clades when using random.clas or BIC/ICL) and, in most instances (n=14/17), they matched the two subspecies: *H. l. hilaris* (12 males) and *H. l. leucophrys* (29 males) (Figures 3.3/S3.6; average trial dendrogram similarity=0.989±0.037). *H. l. hilaris* had one identity song and *H. l. leucophrys* had three (Table S3.10). Repertoires of four F1 hybrid birds clustered with *H. l. leucophrys* while one clustered with *H. l. hilaris*. The six second-generation males clustered according to their parental subspecies and two *H. l. hilaris* males clustered with *H. l. leucophrys*. The identity songs for the two subspecies differ in averaged note peak frequency and minimum/maximum song frequency, with the

*H. l. hilaris* values lower than the *H. l. leucophrys* values. Using random.post as the ECM initialization, two additional genetically *H. l. hilaris* birds clustered with *H. l. leucophrys*. Only the *H. l. hilaris* clade was detected using complete linkage and only the *H. l. leucophrys* clade was detected at the highest value of *critfact*.

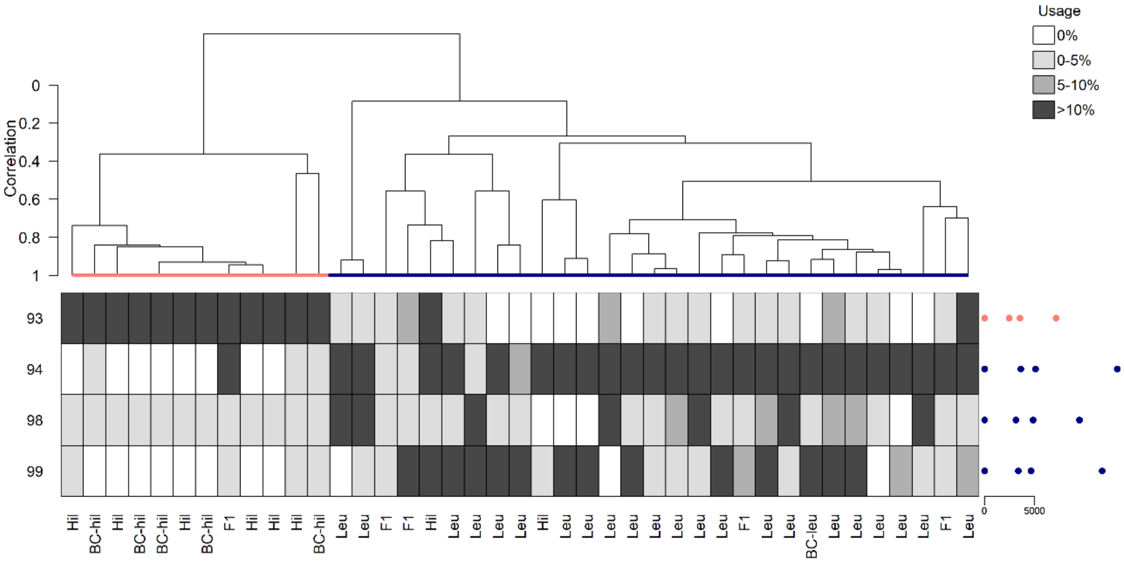


*Figure 3.3* – Baseline dendrogram with subspecies and identity song types for wrens. Average linkage hierarchical clustering dendrogram (top) depicts similarity among song frequency vectors of male wrens. Colored identity clades correspond to two subspecies: *H. l. hilaris* (salmon) and *H. l. leucophrys* (navy). Heat map (bottom) depicts identity song type usage (rows) for each male (columns) in shades of grey, with usage calculated based on probabilistic assignment of songs to types. Identity song codes are on the left and frequency centroid vector plots are on the right. From left to right, the space between the dots represents averaged note peak frequency, minimum song frequency, and maximum song frequency for each song type (scale bar is in Hz). Clustering was done on logged frequency vectors, but vector plots are presented unlogged to aid in interpretation. Genotyping abbreviations are: Hil, parental *H. l. hilaris*; Leu, parental *H. l. leucophrys*; F1, first generation hybrid; BC-hil, backcross between Hil and F1; and BC-leu, backcross between Leu and F1 (Halfwerk et al., 2016).

### *3.4.4 – Crickets*

In most trials (n=16/20), identity clades corresponding to the two species, *T. oceanicus* and *T. commodus*, were detected in the cricket dataset (for baseline dendrogram, see Figures 3.4/S3.7; average trial dendrogram similarity=1.00±0.00). The clade identity song proportion was similar and high for the *T. oceanicus* sites (99%, two identity songs) and the *T. commodus* sites (98%, two identity songs) (Table S3.11). The centroid chirp-trill and trill pulse lengths were shorter in *T. commodus* identity songs than in *T. oceanicus* identity songs. No identity clades were produced at the two highest tested values of *minrep* and only a single identity clade (*T. commodus*) was produced using BIC/ICL.

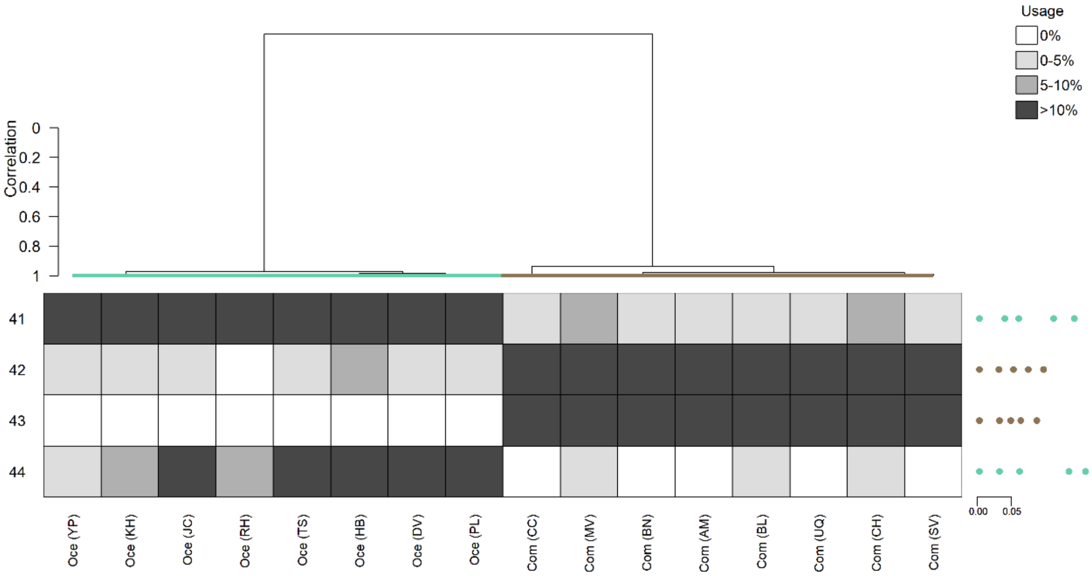


*Figure 3.4* – Baseline dendrogram with species and identity song types for crickets. Average linkage hierarchical clustering dendrogram (top) depicts similarity among song interval vectors of male crickets from 16 sites. Colored identity clades correspond to two species: *Teleogryllus oceanicus* (teal) and *Teleogryllus commodus* (brown). Heat map (bottom) depicts identity song type usage (rows) for each field site (columns) in shades of grey, with usage calculated based on probabilistic assignment of songs to types. Identity song codes are on the left and interval centroid vector plots are on the right. From left to right, the spaces between the dots represent chirp pulse length, chirp interpulse interval, chirp-trill interval, and trill pulse length (scale bar is in seconds). Species abbreviations (Com, *T. commodus*; Oce, *T. oceanicus*) are listed along the bottom, with corresponding field site names in parentheses (Moran et al., 2020).

## 3.4 – Discussion

### *3.5.1 – General method performance*

Contaminated mixture modelling has several strengths as a classification method. It: (1) minimizes the number of parameters specified *a priori*; (2) identifies outliers, letting the user decide if the outliers should be retained (as we did in this study) or excluded; (3) allows clusters to have varying volume, shape, and orientation in multivariate space; and (4) can be used on both low and high dimensional data (Punzo &

McNicholas, 2016). Capitalizing on these advantages, our work demonstrates the versatility of this classification method.

The choice of how to divide calls into repertoires should be guided by features of the user's dataset (e.g. sampling resolution) and research question (e.g. looking for individual vs. group identity signals), as it was here (Method S3.2 in Appendix B). When calls were randomly permuted among repertoires, IDcall did not delineate identity calls or clades in the dendrogram for any dataset for any permutation (n=40 total). Call types that were characteristic of a given repertoire (and may have eventually become identity calls for a clade) were instead scrambled across repertoires. This suggests the method is unlikely to delineate identity calls and clades when they do not exist. In such cases, a dendrogram is still produced but it does not have identity calls or clades. However, the method could also struggle to detect true identity calls and clades for small datasets. For example, when the number of repertoires in the Atlantic/Mediterranean sperm whale dataset was randomly halved, the EC1 (n=31 repertoires) and Mediterranean (n=7) vocal clans were still detected using the default parameters (Figure S3.8) but the EC2 clan (n=3) was not detected until *minrep* was decreased to 3 (Figure S3.9). When the number of repertoires was randomly quartered, no clans were delineated using the default parameters, but the EC1 (n=14) and Mediterranean (n=5) vocal clans were delineated when *minrep* was decreased to 5 (Figure S3.10). The EC2 vocal clan had only one repertoire in this randomized subset, and never became a clan. As these examples illustrate, the method's performance relies on the size and features of the input data.

During call classification, using AIC/AICc generally resulted in the highest total number of call types, ICL the lowest, and BIC intermediate. Using complete linkage during tree building produced the most distinct trees, whereas average and single linkage behaved more similarly. Increasing *critfact* restricted the requirements to be considered an identity call, so fewer calls were identified as such, and increasing *minrep* prohibited small identity clades with few repertoires from forming. Varying the ECM algorithm initialization strategy affected the number of calls (total and identity) and tree identity call proportion, but the direction of these effects varied by dataset. Aside from the wrens (for which random.post misassigned two birds and random.clas did not produce any identity clades), the three strategies generally yielded similar results.

### *3.5.2 – Sperm whales*

Most of the sperm whale clans detected by IDcall agree with past work (Method S3; Tables S3.12/S3.13) and many of the identity codas have previously been recognized as indicators of clan identity (Cantor et al., 2016; Gero et al., 2016; Pavan et al., 2000; Rendell & Whitehead, 2003). Coda type usage results from recent sperm whale studies suggest that the clan identity coda phenomenon extends beyond the locations considered here (Brazil: Amorim et al., 2020; Mauritius: Huijser et al., 2020).

The Atlantic/Mediterranean sperm whale dataset was reliably delineated into three clans across trials except when *minrep* exceeded the number of EC2 clan repertoires. The number and composition of Pacific sperm whale clans was more variable, but this variation centered on the Short clan. The robustness of the other clans (Regular, Plus-One, Four-Plus, and putative fifth) across trials lends confidence to those clan designations. While the repeated segregation of the putative fifth clan from the Four-Plus clan across trials could suggest that IDcall is not performing as well as previous methods (e.g. Cantor et al., 2016), the differences in identity call usage in both identity clades, coupled with no photo-identified groups of whales linking any of the Four-Plus or putative fifth clan repertoires, supports them as separate clans. More data and analyses are necessary, but the current results suggest that IDcall is sensitive enough to identify clans that were not detected with other methods.

The repertoires comprising the baseline dendrogram Short clan typically clustered together across trials but formed one to three clans (Figures 3.2/S3.4/S3.5) depending on the parameters. In the past, the Short clan has been regarded as an anomaly compared to other Pacific clans, given that the clan's most common coda types (e.g. 1+2, 3R, 2+1) do not follow a characteristic rhythmic pattern across click lengths like other clans (Rendell & Whitehead, 2003). Indeed, different combinations of 1+2, 3R, and 2+1 coda types become identity codas for the variations of the Short clan found using IDcall. Short clan coda type usage plots from past work (e.g. Cantor et al., 2016; Rendell & Whitehead, 2003) show that different repertoire sub-groups within the clan vary in how much they use the aforementioned coda types, and each photo-identified group of whales contributed only one repertoire to the present analysis. This suggests that what has previously been referred to as the Short clan could be one or several clans, but more acoustic and photographic data are needed to tease these possibilities apart.

### *3.5.3 – Wrens*

Using frequency measures, IDcall accurately clustered songs from male wrens into two subspecies. Consistent with the original study (Halfwerk et al., 2016), two *H. l. hilaris* males clustered with *H. l. leucophrys* in the baseline dendrogram. The distribution of F1 hybrid birds in the baseline dendrogram aligns with the original study as well (Halfwerk et al., 2016). The identity songs were consistent across trials and emphasized known song difference (Method S3; Halfwerk et al., 2016). That only the *H. l. leucophrys* clade was detected at the highest level of *critfact* suggests that *H. l. leucophrys* birds use their 'identity songs' more than *H. l. hilaris* birds.

### *3.5.4 – Crickets*

The cricket dataset was robust to parameter variation. In most trials, IDcall accurately clustered male cricket calling songs into two species, and the detected identity songs emphasized known song differences (Method S3; Moran et al., 2020). Unsurprisingly, the two species were not delineated when *minrep* was increased beyond the number of repertoires for each species (n=8). While no wren identity clades and only one of the two expected cricket identity clades (*T. commodus*) were detected when using BIC/ICL, both datasets are small and AICc is likely the most appropriate choice for information criterion (Hurvich & Tsai, 1989).

### *3.5.5 – Method recommendations, limitations, and applications*

The IDcall options/parameters that were varied during our trials can be set by the user and informed by features of their dataset and research question. Using *k*-means as the ECM initialization strategy is effective when there are numerous local optima, and likely represents a good starting point for most users, but random.post and random.clas can be more computationally efficient (Shireman et al., 2017). Prior to call classification, users should examine their raw call data to get a better sense for how separated call clusters are. The number of mixture components (which BIC aims to find) may not always equal the number of clusters (which ICL aims to find), which can lead to different results depending on the dataset (Baudry et al., 2010). Generally, we recommend using BIC for large datasets with poorly-separated clusters (Schwarz, 1978); ICL for large datasets with well-separated clusters (Biernacki et al., 2000); and AICc for small datasets (Hurvich & Tsai, 1989). Users can also set the range of number of mixture components to fit to the data during call classification.

Average linkage represents a middle ground between the extremes of single linkage and complete linkage and is likely an appropriate linkage method for most datasets. Using high values of *critfact* will generally result in fewer but more conservative identity calls (i.e. identity calls that are used very frequently by their identity clade and very rarely by all other clades) and using high values of *minrep* will typically result in fewer, larger identity clades. Users can thus vary *critfact* and *minrep* based on their desired level of conservativeness at both the identity call and identity clade level. The comprehensive output also lets users control metric cut-offs like repertoire/clade correlation. Repertoires that have low call type usage correlation with the rest of an identity clade can be excluded from subsequent analyses.

IDcall does have limitations. It will not cluster calls represented by different vector lengths (i.e. occupying multivariate spaces with varying dimensionality, *N*). It also requires several parameters to be set, and these settings can impact the final dendrograms, as evidenced by some of the trial results. We recommend that users test different parameter combinations and compare the dendrograms. Dendrogram regions that are more sensitive warrant careful interpretation but may reflect underlying uncertainty in the data and/or true biological variability, whereas regions that are robust can likely be interpreted with more confidence.

We make no claims that IDcall will work for all acoustic animals, but its ability to find and use identity calls to detect structure—including subpopulations, subspecies, and species—in three diverse taxa is promising. However, the presence of biologically informative identity calls does not necessarily mean those calls are meaningful to the animals themselves. Playback experiments are key to test whether specific calls are used by animals to broadcast or infer identity. IDcall can inform such studies by identifying calls that are potentially characteristic of underlying biological structure (and likely to be perceptually salient) and are thus good targets for additional research or to be played back. For animals that are cryptic or difficult to access but have identity calls, IDcall can potentially aid managers in quantifying behavioral diversity and/or identifying putative discrete units, given that acoustic data can be inexpensive and minimally invasive to collect. This is particularly true in marine environments, where passive acoustic recording is relatively easy whereas genetic or photographic sampling is often logistically complex.

## 3.5 – Acknowledgements

We thank: Mauricio Cantor, Lindy Weilgart, Anne Bøttcher, and Ellen Jacobs for annotating codas; Kristian Beedholm for creating and sharing Coda Sorter; Noor Youssef for offering method advice; Scott McCain and the Stack Overflow community for answering coding questions; Jack Rayner for directing us to the cricket dataset; Wouter Halfwerk for answering questions about the wren dataset; Laela Sayigh, Gianni Pavan, Sara Rose, Mary Ann Daher, and Kurt Fristrup for helping with the Watkins Marine Mammal Sound Database; and Paul McNicholas for advising on contaminated mixture models. We are grateful to the individuals and organizations that enabled the wren and cricket research and to the authors for making the data accessible.

# CHAPTER 4

# CULTURAL STRUCTURING AND SYMBOLIC MARKING IN SPERM WHALES ACROSS THE PACIFIC OCEAN[17]

> "Sometimes you just find a culture that breaks your heart."
> —Lily King

## 4.1 - Abstract

Culture is increasingly being recognized as a powerful force in structuring human and non-human animal populations. For sperm whales, culture segments populations into vocal clans with distinctive phenotypes, including clan-specific dialects comprised of stereotyped click patterns, called codas. We analyzed 23,429 codas from 23 locations to characterize sperm whale culture across the Pacific Ocean. Repertoires of codas were delineated into seven clans using the identity call method, in which codas are classified into types using contaminated mixture models and repertoires are hierarchically clustered into clans based on coda type usage. We then investigated how acoustic similarity (measured as similarity in usage of different coda types) varied with geographic distance within clans or with spatial overlap between clans. Our results suggest that both drift and selection are at play in Pacific sperm whale dialects, with different coda types undergoing different processes. Within-clan usage of 'identity codas'—coda types that are diagnostic of clan identity—is more stable over geographic space than non-identity coda usage, and between-clan acoustic similarity calculated using identity codas decreases with increasing spatial overlap, whereas non-identity coda usage does not change. Together, these results

---

[17] This chapter is in collaboration with: Shane Gero, Luke Rendell, Lindy Weilgart, Maurício Cantor, Masao Amano, Elisabeth Slooten, Chris Johnson, Iain Kerr, Roger Payne, Andy Rogan, Ricardo Antunes, Olive Andrews, Elizabeth Ferguson, Cory Ann Hom-Weaver, Thomas Norris, Yvonne Barkley, Karlina Merkens, Erin Oleson, Thomas Doniol-Valcroze, James Pilkington, Jonathan Gordon, Manuel Fernandes, Marta Guerra, Leigh Hickmott, and Hal Whitehead. SG conceived the project, TAH analyzed and interpreted the data, and HW helped conceive the methods. SG and HW provided input throughout and reviewed drafts of the chapter. All other collaborators contributed sperm whale recordings and/or codas.

provide empirical evidence that identity codas function as symbolic markers of sperm whale cultural identity, like ethnic markers in humans. This first ‘cultural census’ of Pacific sperm whales provides baseline information that can facilitate culturally-sensitive conservation and management of this deep-sea keystone species.

## 4.2 – Introduction

Culture—defined here as information or behavior that is shared within a community and acquired from conspecifics through social learning—can structure populations in a variety of ways (Whitehead & Rendell, 2014). This is abundantly clear for humans (*Homo sapiens,* Bell et al., 2009), but mounting evidence shows that it applies to other animals as well (reviewed in Allen, 2019). For sperm whales (*Physeter macrocephalus*), culture is a pervasive aspect of life (reviewed in Whitehead, 2003).

In social situations, sperm whales communicate using stereotyped patterns of clicks, called codas, which can be divided into types based on the number of clicks and inter-click intervals (ICIs). Even in sympatry, whales only associate with other individuals that use similar coda types (i.e. have a similar dialect), forming a cultural level of population structure called the vocal clan (Rendell & Whitehead, 2003b). There is no indication that clans differ in the distribution of nuclear genes (Konrad et al., 2018; Whitehead, 2003), although there is similarity among clan members in mitochondrial haplotypes, presumably due to consistent clan membership within matrilines (Rendell et al., 2012). The xenophobia seen in areas of clan sympatry suggests that sperm whales can discriminate cultural in-group vs. out-group members. In humans, such discrimination is facilitated by symbolic markers—seemingly arbitrary traits that indicate cultural group membership, such as dialects, clothing, or rituals (Boyd & Richerson, 1987; Cohen, 2012; McElreath et al., 2003). Given that sperm whale clan dialects appear to be the product of cultural transmission via biased social learning (Cantor et al., 2015), it has been posited that they may act as symbolic markers of clan identity (Cantor & Whitehead, 2013; Gero, Whitehead, et al., 2016; Hersh et al., 2021; Rendell & Whitehead, 2003b), but empirical evidence is lacking. However, recent work by Hersh et al. (2021) found that sperm whale clans in both the Atlantic and Pacific could be distinguished using specific coda types, which they termed ‘identity codas’. If these identity codas can be used by researchers to tell different clans apart, might they serve a similar function for the whales themselves?

In the Pacific Ocean, sperm whale coda research has primarily been conducted off South America (Rendell & Whitehead, 2003b) and Japan (Amano et al., 2014). Interesting patterns have emerged from these efforts, including the existence of sympatric clans and rhythmic motifs in some clan dialects (Amano et al., 2014; Cantor et al., 2016; Rendell & Whitehead, 2003b). The clans identified by Rendell & Whitehead (2003b) are thought to contain thousands of individuals and to have large ranges as a whole, although the approximate annual home range span of individual eastern tropical Pacific sperm whales is just 1,000 km (Whitehead, 2001; Whitehead et al., 2008). This suggests that whales belonging to a single clan are unlikely to ever meet all other members of that clan. It is unknown whether the trends documented off South America and Japan are region-specific or extend to other Pacific sperm whales. The overarching goal of the present research was to quantify sperm whale cultural diversity on an ocean-basin scale. This goal was made tractable through collaboration with many researchers and organizations from across the Pacific, enabling the largest-scale sperm whale coda study to date. Our objectives were threefold:

1) Quantify the number and spatial extent of clans in the Pacific Ocean.
2) Determine how acoustic similarity[18] varies with spatial distance within clans.
3) Determine how acoustic similarity varies with spatial overlap between clans.

The first objective is essentially a 'cultural census' of Pacific sperm whales, which are classified as vulnerable in the International Union for Conservation of Nature Red-List. The conservation implications of animal culture are increasingly being recognized, particularly for sperm whales (Brakes et al., 2019, 2021; Convention on Migratory Species, 2017), but a key first step towards incorporating culture into management is understanding modern levels of cultural diversity. The second objective

---

[18] Acoustic similarity was measured as similarity in usage of different coda types (e.g. identity coda types, non-identity coda types, all coda types) within and between clan dialects. We use the term acoustic similarity (as opposed to, for example, coda type usage similarity) for brevity, but note that it was always calculated in terms of coda type usage.

aims to discern which evolutionary processes are at play in the identity and non-identity codas comprising dialects. The final objective addresses whether identity codas are used as symbolic markers of clan identity by quantifying if and how identity and non-identity coda usage are modulated by clan overlap (a proxy for degree of sympatry). If identity codas are used as symbolic markers of clan identity, we hypothesize that their usage within clans will be more stable over geographic distance (due to selection) than non-identity coda usage, and that their usage will become more distinct as clan overlap increases, while no change is predicted for non-identity coda usage.

## 4.3 – Methods

### *4.3.1 – Data collection and coda extraction*

Acoustic recordings of sperm whales were collected between 1978 and 2017 in 23 regions spanning 93 degrees of latitude and 154 degrees of longitude in the Pacific Ocean (for region abbreviations and methods, see Table S4.1). For 12 regions (starred in Table S4.1), codas had previously been extracted and described (in terms of number of clicks and ICIs) (Amano et al., 2014; Cantor et al., 2016; Hersh et al., 2021; Rendell & Whitehead, 2003b; Weilgart & Whitehead, 1997), although additional codas were sometimes extracted for those regions in the present study as well. TON codas were extracted but not yet published. For the remaining regions, raw audio recordings of sperm whales were audited for codas in Audacity (version 2.3.0)[19]. Coda parameters were extracted using 'Coda Sorter'—a custom software implemented in LabView and run in MATLAB (version R 2020a). Extracted codas with 3–10 clicks were divided into repertoires by recording day (such that each repertoire is comprised of all the codas recorded in a specific region on a single day; Table S4.2). Repertoire locations were inferred from satellite navigation data (with varying levels of precision; see Table S4.1). When multiple fixes were available for a given day of recordings (i.e. for a given repertoire), we took the average as the repertoire location. For repertoires with identical recording locations, we jittered the locations using the 'geoR' R package (Ribeiro Jr. & Diggle, 2001) prior to analyses.

---

[19] Audacity® software is copyright © 1999-2021 Audacity Team. The name Audacity® is a registered trademark.

### *4.3.2 – Coda typing and clan assignment*

IDcall (Hersh et al., 2021) was used to classify codas into types and assign repertoires to clans. In IDcall, codas are first indexed in multivariate space using ICIs. Codas with the same number of ICIs (i.e. the same click length) are then classified into types using parsimonious mixtures of multivariate contaminated normal distributions, with each coda assigned a probability of belonging to each type. Repertoires are hierarchically clustered based on similarities in coda type usage. An ‘identity clade’ (i.e. putative sperm whale clan) of repertoires is denoted if it has at least one ‘identity coda’ (i.e. a coda type made frequently in that clade but rarely in any other). The output includes a dendrogram of repertoires colored by putative clan, a plot showing the rhythmic click pattern of the clans’ identity coda(s), and a heat map of identity coda usage across clans. Using the IDcall ‘postprob’ extension (Whitehead & Hersh, submitted), we also calculated the posterior probabilities (using all coda types) that each repertoire is a member of each identity clade

We used the default parameter settings from Hersh et al. (2021) with a few deviations (see Method S4.1 in Appendix C, Figure S4.1). With *minrep* (i.e. the minimum number of repertoires required for an identity clade to form) conservatively set to 15, we tested a range of values for the *critfact* parameter (i.e. the factor by which a coda type must be made more in a given clade compared to other clades for it to become an identity coda) during hierarchical clustering and compared the resultant dendrograms to determine the most probable number of sperm whale clans in the Pacific Ocean. Maps showing clan distributions were created using the ‘rgeos’ (Bivand et al., 2017), ‘sf’ (Pebesma, 2018), and ‘ggspatial’ R packages

### *4.3.3 – Within-clan analyses*

Within each clan, our goal was to determine whether acoustic similarity of repertoires varied with geographic distance between repertoires. For each pair of repertoires within a clan, we plotted the logged geographic distance separating them (see Method S4.2 in Appendix C) against their acoustic similarity and fit a linear regression using the ‘stats’ R package. The regression slope indicated if and how acoustic similarity varied with increasing distance. In IDcall, hierarchical clustering is based on the correlation of coda type usages between repertoire pairs (Hersh et al., 2021). In the present study, the acoustic similarity of each pair of repertoires within clans was calculated based on correlations in usage of identity coda types ($acsim_{IDwi}$) or non-identity coda types ($acsim_{nonIDwi}$), and the results were compared. We ran one-tailed Mantel tests with 10,000 permutations to assess the direction and significance of any correlation. All Mantel tests used the 'ade4' R package (Dray & Dufour, 2007) and significance was set at $p \leq 0.05$.

Having found a general pattern of decreasing acoustic similarity with increasing geographic distance, we tested the null hypothesis that acoustic similarity of repertoires within a clan, calculated using just identity codas or just non-identity codas, drops off similarly with increasing distance due to drift. However, if identity codas are used by sperm whales as symbolic markers of clan identity, we predicted that within-clan identity coda usage would be more stable over geographic space than non-identity coda usage, with non-identity codas more susceptible to drift. We explicitly tested these hypotheses by plotting the difference in $acsim_{IDwi}$ and $acsim_{nonIDwi}$ (hereafter $acsim_{diffwi}$) against logged geographic distance and used one-tailed Mantel tests with 10,000 permutations to assess the direction and significance of any correlation within each clan. We then combined the p-values from different clans using Stouffer's method weighted by the square root of the number of repertoires in each clan to test for overall trends (Heard & Rubin-Delanchy, 2018).

### *4.3.4 – Between-clan analyses*

Between clans, our goal was to determine whether acoustic similarity varied with spatial overlap. For each pair of clans, spatial overlap was calculated as the proportion of the first clan's repertoires that were recorded within 1,000 km (approximate annual home range span of eastern tropical Pacific sperm whales; Whitehead, 2001; Whitehead et al., 2008) of at least one of the second clan's repertoires. This measure of spatial overlap does not consider temporal overlap; given the patchy sampling of codas in different regions in different years (with 61% of regions represented by one year and 87% represented by one to two years; Table S4.2), our analyses do not include a temporal component. This means that, theoretically, two clans could have a non-zero spatial overlap value without having been recorded in the same year.

Between-clan acoustic similarity was calculated by first averaging the identity or non-identity coda type usages across the repertoires in each clan into a single repertoire per clan, and then calculating the overall correlation in coda type (identity or non-identity) usage for all clan pairs. We then plotted the minimum, mean, and maximum spatial overlap against $\text{acsim}_{\text{IDbt}}$ or $\text{acsim}_{\text{nonIDbt}}$ for all clan pairs. We used one-tailed Mantel tests with 10,000 permutations to determine the direction and significance of any correlation, testing the null hypothesis that acoustic similarity between clans does not decline with increasing spatial overlap.

Having observed a general pattern of decreasing $\text{acsim}_{\text{IDbt}}$ (but not $\text{acsim}_{\text{nonIDbt}}$) with increasing geographic distance, we tested the null hypothesis that between-clan acoustic similarity, whether calculated as $\text{acsim}_{\text{IDbt}}$ or $\text{acsim}_{\text{nonIDbt}}$, falls off similarly with increasing clan overlap. If identity codas are used by sperm whales in an identifying role, we predicted that $\text{acsim}_{\text{IDbt}}$ would decrease with increasing clan overlap, because whales would modulate identity coda production in some way that enhances clan distinctiveness in areas of greater overlap. In contrast, we did not expect to see a trend in $\text{acsim}_{\text{nonIDbt}}$ if, as we posit, non-identity codas are generally not used as markers of clan identity. We explicitly tested this hypothesis by plotting $\text{acsim}_{\text{diffbt}}$ against spatial overlap and used one-tailed Mantel tests with 10,000 permutations to assess the direction and significance of any correlation across clans.

### *4.3.5 – Alternative analyses*

We carried out a range of additional analyses using alternative acoustic similarity measures (e.g. based on correlations in usage of all coda types (i.e. both identity and non-identity) within ($\text{acsim}_{\text{allwi}}$) and between ($\text{acsim}_{\text{allbt}}$) clans) and methods (e.g. multidimensional scaling to visualize clan acoustic similarity) (see Appendix C) and note when any of these results substantially differ from those presented in the main text.

## 4.4 – Results

### *4.4.1 – Coda dataset*

In total, 23,429 codas with 3–10 clicks were extracted from Pacific Ocean sperm whale recordings (Figure 4.1) and classified into types (Table S4.2). Shorter and longer codas (n=808) were excluded because they made up only 4.35±8.81%[20] (mean±SD) of each regional dataset and have been inconsistently marked across studies. Only well-sampled repertoires (i.e. with at least 25 codas, Rendell & Whitehead, 2003a) were included in the clan analyses (22,829 codas, 191 repertoires).

---

[20] This decreases to 2.68±3.79% if the codas from MNP are excluded. One hundred of the 278 MNP codas were 2-click codas.

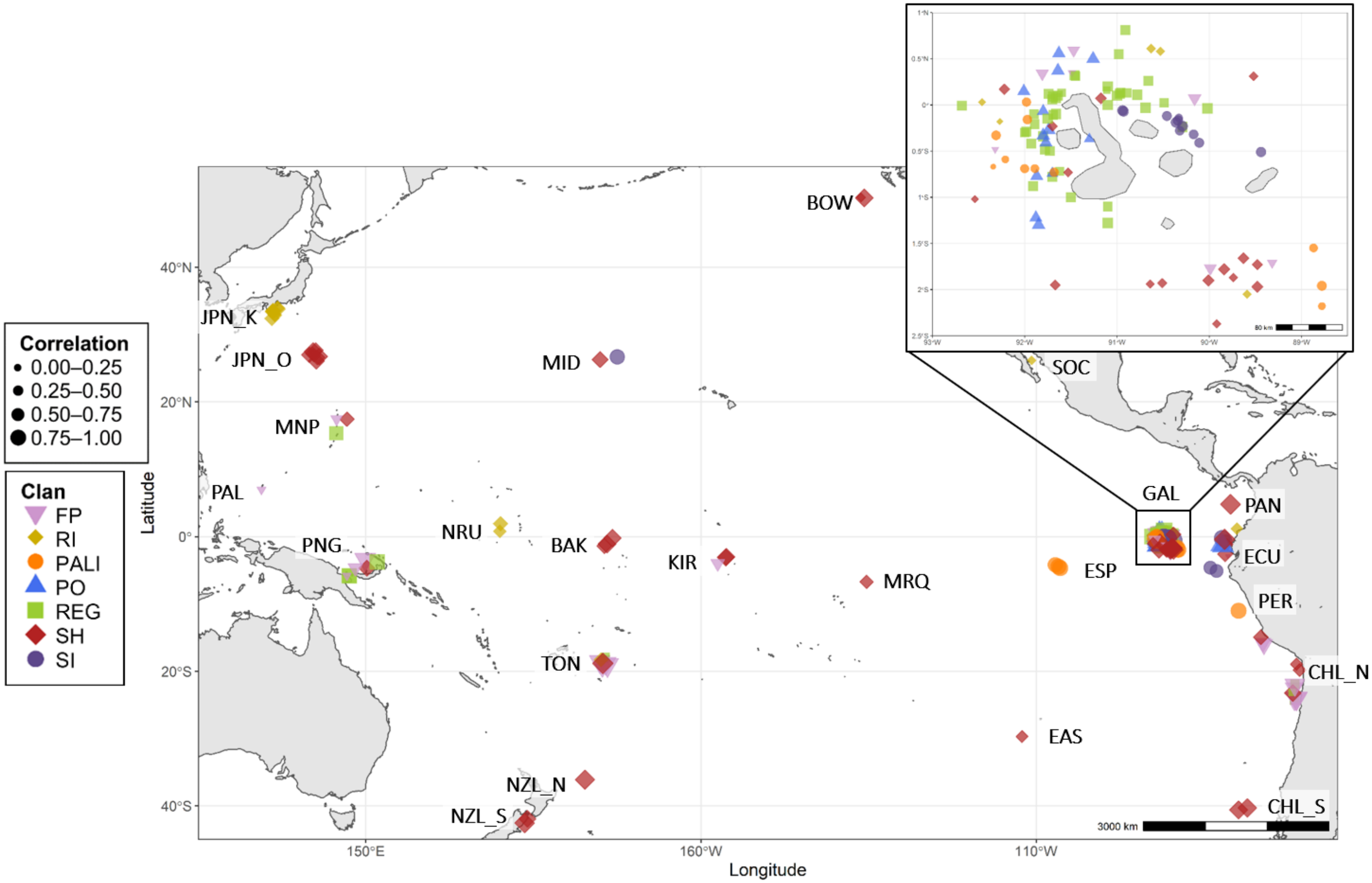


*Figure 4.1* – Composite map of sperm whale clan distribution across 23 regions in the Pacific using the 7-clan tree clan designations. See Table S4.1 for region abbreviations. Each shape is a single repertoire, colored according to clan. The size of each shape indicates the within-clan correlation of that repertoire, with smaller shapes having lower correlations. The scale bar is approximate and most accurate along the equator. Clan abbreviations are FP=Four-Plus, RI=Rapid Increasing, PALI=Palindrome, PO=Plus-One, REG=Regular, SH=Short, and SI=Slow Increasing.

### *4.4.2 – Clans and identity codas*

Varying *critfact* from 3 to 20 while keeping *minrep* constant at 15 produced three reasonable (i.e. at least 75% of repertoires were assigned to a clan) scenarios: a 5-clan tree (Figures S4.2/S4.3), a 7-clan tree (Figures 4.2/S4.4), and an 8-clan tree (Figures S4.5/S4.6) (Table S4.3). The tree cophenetic correlation coefficient is 0.905 and four clans are robustly delineated: three well-known clans (the 'Four-Plus', 'Plus-One', and 'Regular' clans; Rendell & Whitehead 2003b; Cantor et al. 2016) and a clan first described as putative in Hersh et al. (2021) using a much smaller and geographically restricted dataset. As in Hersh et al. (2021), variation across trees centered on repertoires dominated by shorter (mainly 3- and 4-click) codas, with these repertoires (n=85) forming one, three, or four clans, respectively, in the aforementioned trees. The available evidence supports the 7-clan tree (with three of these 'shorter coda' clans; see Discussion S4.1 in Appendix C) and will be referenced hereafter but note that the division of repertoires dominated by shorter codas into clans is more uncertain than the division of repertoires into the four robust clans. This is clearly illustrated by the posterior probability plots in Figures 4.2 (7-clan tree), S4.2 (5-clan tree), and S4.5 (8-clan tree).

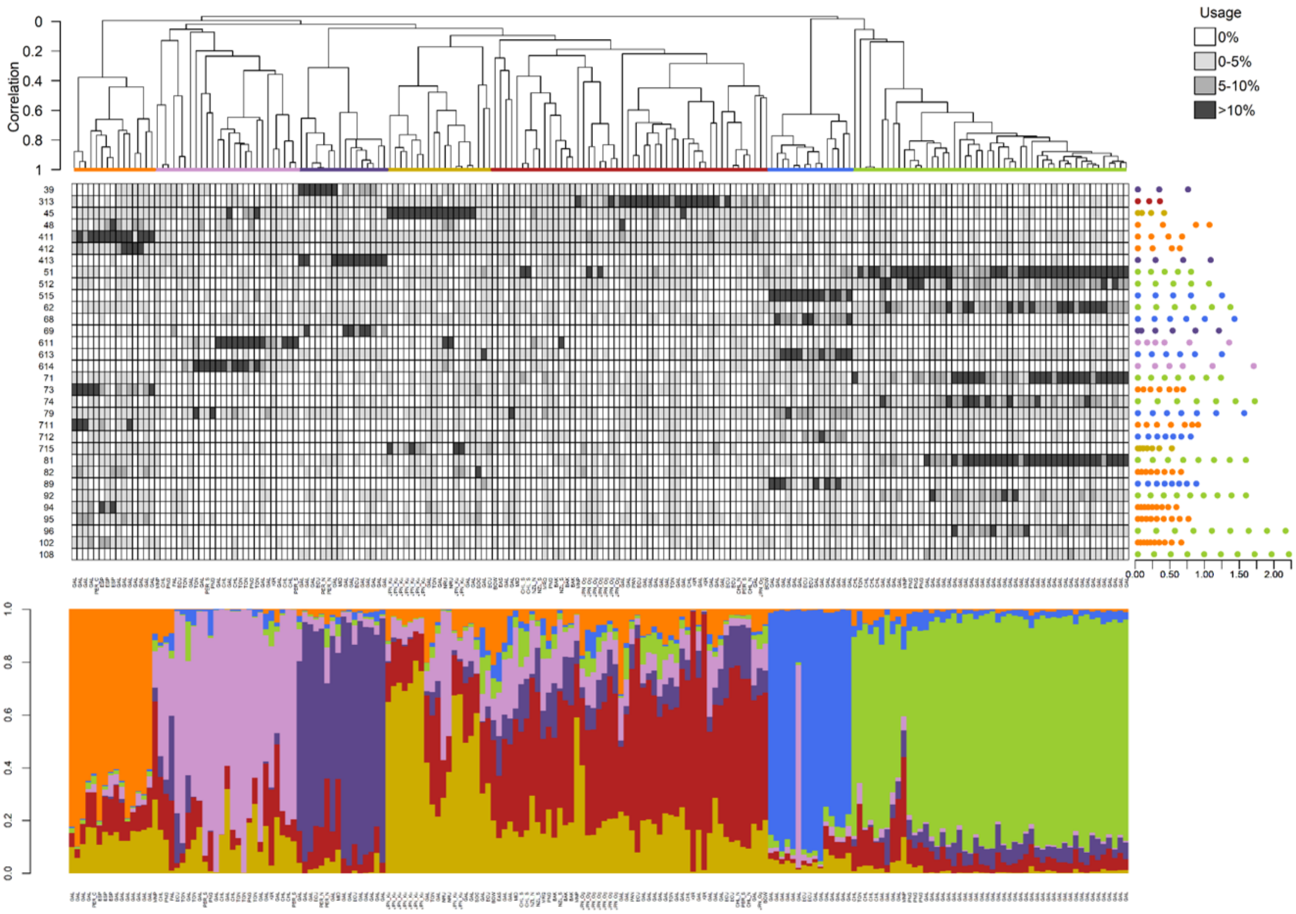


*Figure 4.2* – Seven-clan tree with identity coda types. This tree was constructed using *critfact*=5 and *minrep*=15 (Table S4.3). Average linkage hierarchical clustering dendrogram (top) depicts similarity among sperm whale coda repertoires recorded across the Pacific Ocean. Heat map (middle) depicts identity coda type usage (rows) for each repertoire (columns) in shades of grey. Usage is calculated based on probabilistic assignment of codas to types. Identity coda type codes are on the left (see Table 4.1 for type names) and type centroid rhythm plots (colored by clan; see Figure 4.1) are to the right of the heat map (each dot represents a click; scale bar is in seconds). Stacked bar plot (bottom) shows the posterior probabilities of repertoire assignations to clans (Whitehead & Hersh, submitted). Recording locations are listed along the bottom (see Table S4.1 for abbreviations).

Of the seven clans, one is 'new'[21] and two were previously documented but not named (Amano et al., 2014; Hersh et al., 2021). We refer to the clan first described in Hersh et al. (2021) as the 'Palindrome' clan, in homage to the palindromic nature of the clan's most frequently used identity codas (Figure 4.2, in orange). We suggest referring to the clan first described by Amano et al. (2014) off the Kumano coast as the 'Rapid Increasing' clan, given that their work and ours show that very short duration codas with increasing ICIs predominate in this clan (Figure 4.2, in gold). Finally, the 'new' clan also makes codas with increasing ICIs, but the overall duration of these codas is much longer than in the Rapid Increasing clan; we suggest referring to this clan as the 'Slow Increasing' clan (Figure 4.2, in purple).

The number of repertoires and within-clan repertoire correlation varies across clans, with some clans showing high consistency in coda type usage across repertoires (e.g. Plus-One, Regular) and others showing much lower consistency (e.g. Four-Plus, Short) (Table 4.1; for 5-clan results, see Table S4.4; for 8-clan results, see Table S4.5). The number of identity codas per clan also varies (Table 4.1), and most clans display rhythmic 'motifs' (Rendell & Whitehead, 2003b) in their identity codas (Figure 4.2). In line with past work, we found that Regular clan whales frequently make codas with equally spaced (i.e. isochronous) clicks, Plus-One clan whales make codas with an extended pause before the final click, and Four-Plus clan whales make codas with a root of four isochronous clicks (Rendell & Whitehead, 2003b). The Palindrome, Rapid Increasing, and Short Increasing clans also appear to have rhythmic motifs in their identity codas (as described in the previous paragraph). Codas from Short clan whales do not seem to follow a rhythmic motif but are typically comprised of very few clicks (for example, the single Short clan identity coda has three clicks) (Figures 4.2/S4.4). Across clans, 3–6-click codas dominated, with exact preferences varying by clan (Figure S4.7).

---

[21] By new, we mean unaccounted for in the published literature.

*Table 4.1* – Summary of clans and identity codas in the 7-clan tree. For each repertoire, we calculated the correlation between the coda type usages of the repertoire and the median usages of the clan; the within-clan correlation was calculated by averaging these values. For each identity coda type, we list the number of clicks, the numeric code (see Figure 4.2), and the type name (following conventions in Weilgart and Whitehead 1997; Hersh et al. 2021). Briefly, type names refer to the overarching rhythmic pattern of clicks in each coda regardless of total duration, which is why multiple numeric codes can be linked to one type. 'R' stands for 'Regular' (i.e. all ICIs are approximately equal), '+' denotes an extended pause between clicks, 'D' stands for 'Decreasing' (i.e. ICIs become shorter throughout the coda), and 'I' stands for 'Increasing' (i.e. ICIs become longer throughout the coda). See Figure 4.1 for clan abbreviations.

| **Clan** | **Number of repertoires** | **Within-clan correlation (mean±SD)** | **Number of identity codas** | **Identity coda types** | | |
|---|---|---|---|---|---|---|
| | | | | *Clicks* | *Numeric code* | *Type names* |
| PALI | 15 | 0.699±0.197 | 9 | 4 | 48, 411, 412 | 1+1++2, 2+2, 2+2 |
| | | | | 7 | 73, 711 | 3+1+3, 3+1+3 |
| | | | | 8 | 82 | 8I |
| | | | | 9 | 94, 95 | 9I, 9I |
| | | | | 10 | 102 | 10I |
| FP | 26 | 0.445±0.314 | 2 | 6 | 611, 614 | 4+1++1, 4+1++1 |
| SI | 16 | 0.730±0.175 | 3 | 3 | 39 | 2+1 |
| | | | | 4 | 413 | 4I |
| | | | | 6 | 69 | 6I |
| RI | 19 | 0.652±0.185 | 2 | 4 | 45 | 4I |
| | | | | 7 | 715 | 7I |
| SH | 50 | 0.484±0.205 | 1 | 3 | 313 | 3R |
| PO | 15 | 0.854±0.123 | 6 | 5 | 515 | 4+1 |
| | | | | 6 | 68, 613 | 5+1, 5+1 |
| | | | | 7 | 79, 712 | 6+1, 7R |
| | | | | 8 | 89 | 8R |
| REG | 50 | 0.763±0.243 | 9 | 5 | 51, 512 | 5R, 5R |
| | | | | 6 | 62 | 6R |
| | | | | 7 | 71, 74 | 7R, 7R |
| | | | | 8 | 81 | 8R |
| | | | | 9 | 92, 96 | 9R, 9R |
| | | | | 10 | 108 | 10R |

### *4.4.3 – Clan distribution*

Clan distribution across the Pacific varied, with some clans showing small ranges while others spanned the ocean basin (Figures 4.1/S4.8). Given the unequal sampling across regions, these maps provide positive evidence of where clans have been detected but should be used cautiously to assess clan absence, especially for regions with few repertoire recordings. The Plus-One clan has only been detected off GAL and ECU so far, while the Short clan has been detected (with varying levels of certainty) in 18 of the 23 regions sampled here. The Palindrome and Slow Increasing clans seem to be primarily detected in South American waters, while the Four-Plus, Rapid Increasing, and Regular clans are more widespread. Geographically sympatric clans were detected in 9/23 Pacific regions, with the highest diversity detected off GAL (7 clans), ECU (5 clans), TON (4 clans), and CHL_N (4 clans). There is a general decrease in the number of detected clans per region as the number of repertoires and the number of codas decreases (especially the former; Table S4.6).

### *4.4.4 – Within-clan trends*

Plots of correlations in usage of identity codas between pairs of repertoires within a clan ($acsim_{IDwi}$) against logged geographic distance have a negative slope for all clans except Plus-One (Figure 4.3). This trend was significant for some (Rapid Increasing, Short, and Regular) but not all (Palindrome, Four-Plus, and Slow Increasing) of the clans (Figure 4.3). The Plus-One clan had a slightly positive but non-significant slope. In general, a decrease in similarity becomes apparent at ranges greater than 1,000 km (the approximate span of the Plus-One clan).

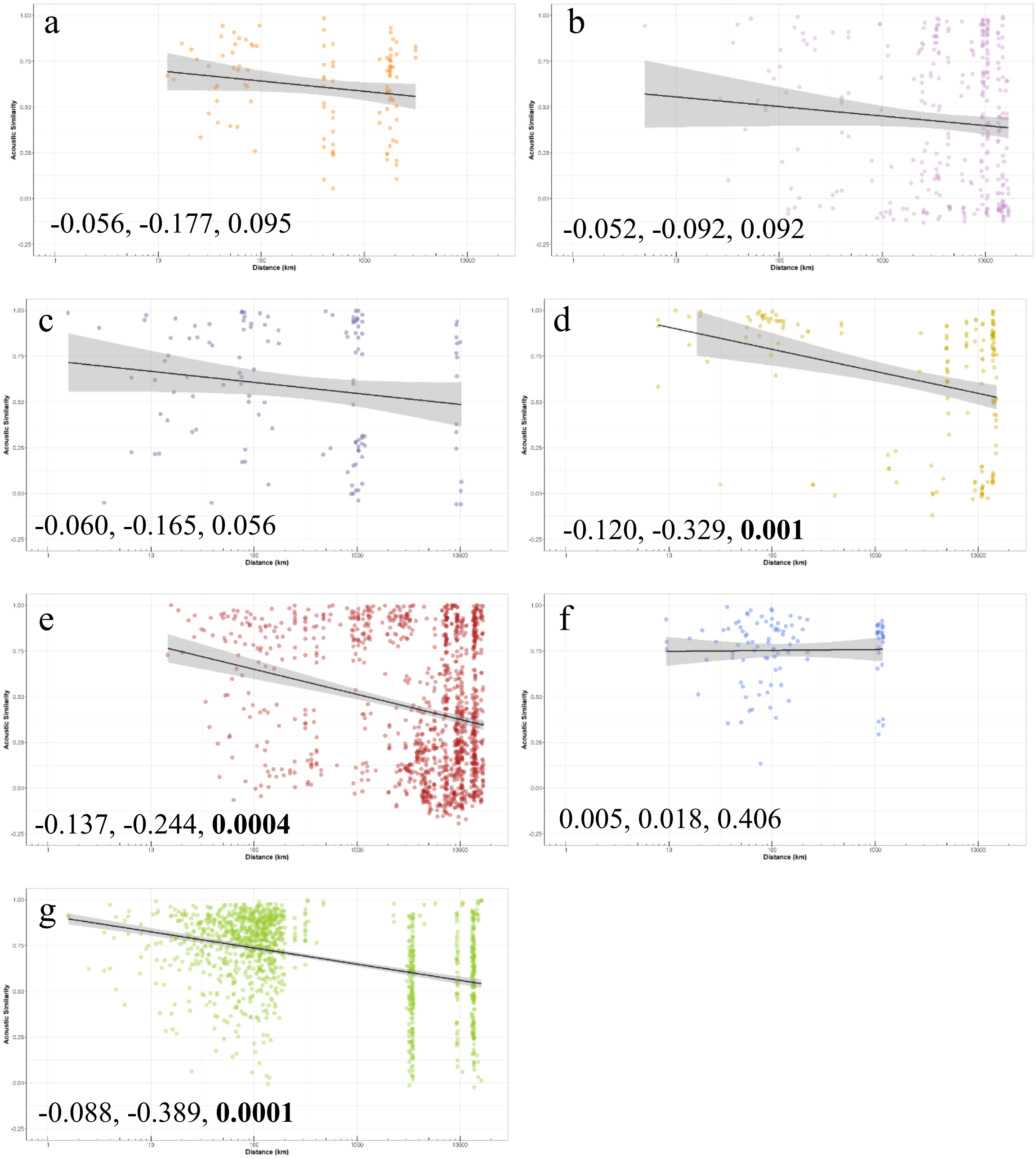


*Figure 4.3* – Logged geographic distance vs. $acsim_{IDwi}$. The 95% confidence intervals are shown in gray. Each dot represents a pair of repertoires. For each panel, the regression line slope, Mantel test matrix correlation, and Mantel test p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that $acsim_{IDwi}$ decreases as geographic distance increases. Panels correspond to the: (a) Palindrome; (b) Four-Plus; (c) Slow Increasing; (d) Rapid Increasing; (e) Short; (f) Plus-One; and (g) Regular clans.

Plots of correlations in usage of non-identity codas between pairs of repertoires within clans ($acsim_{nonIDwi}$) against logged geographic distance also showed a negative slope for all clans (Figure 4.4). This trend was significant for most clans (Palindrome, Four-Plus, Rapid Increasing, Short, and Regular). For 5/7 clans, the slope of the linear regression was steeper when using $acsim_{nonIDwi}$ than when using $acsim_{IDwi}$ (Figures 4.3/4.4), providing support for our alternative hypothesis that within-clan identity coda usage is more stable over geographic space than non-identity coda usage (see Figure S4.9 for the results using $acsim_{allwi}$). This pattern was reiterated in our analysis of geographic distance vs. $acsim_{diffwi}$, with most plots having a positive linear regression slope (because $acsim_{nonIDwi}$ decreases more with increasing distance, making the difference between $acsim_{IDwi}$ and $acsim_{nonIDwi}$ larger at greater distances; Figure S4.10). The trend was significant for the Palindrome and Rapid Increasing clans but was not significant for all clans combined (Stouffer weighted p=0.180). Thus, we could not reject the null hypothesis that, across all clans, acoustic similarity of repertoires within a clan, calculated using just identity codas or just non-identity codas, drops off similarly. However, the results for most clans individually show this trend.

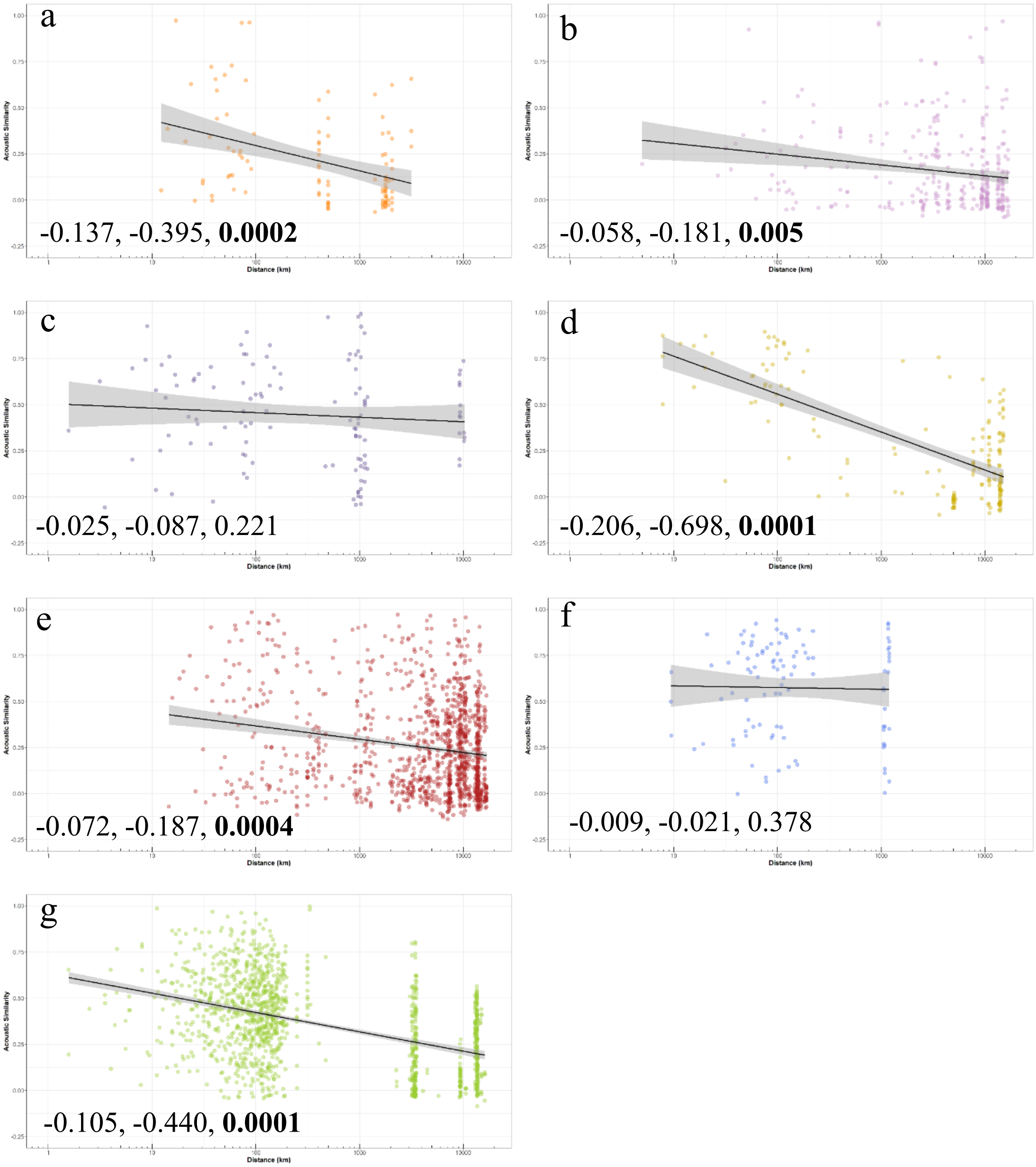


*Figure 4.4* – Logged geographic distance vs. $\text{acsim}_{\text{nonIDwi}}$. The 95% confidence intervals are in shown in gray. Each dot represents a pair of repertoires. For each panel, the regression line slope, Mantel test matrix correlation, and Mantel test p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that $\text{acsim}_{\text{nonIDwi}}$ decreases as geographic distance increases. Panels correspond to the: (a) Palindrome; (b) Four-Plus; (c) Slow Increasing; (d) Rapid Increasing; (e) Short; (f) Plus-One; and (g) Regular clans.

*4.4.5 – Between-clan trends*

Plots of correlations in usage of identity codas between pairs of clans ($acsim_{IDbt}$) vs. the spatial overlap of the clans (Table S4.7) had a negative slope, and Mantel tests confirmed that this trend was significant for mean (p=0.043) and maximum (p=0.048) overlap and marginally significant (p=0.073) for minimum overlap (Figure 4.5). Identity coda usage similarity between clan pairs decreases as spatial overlap increases.

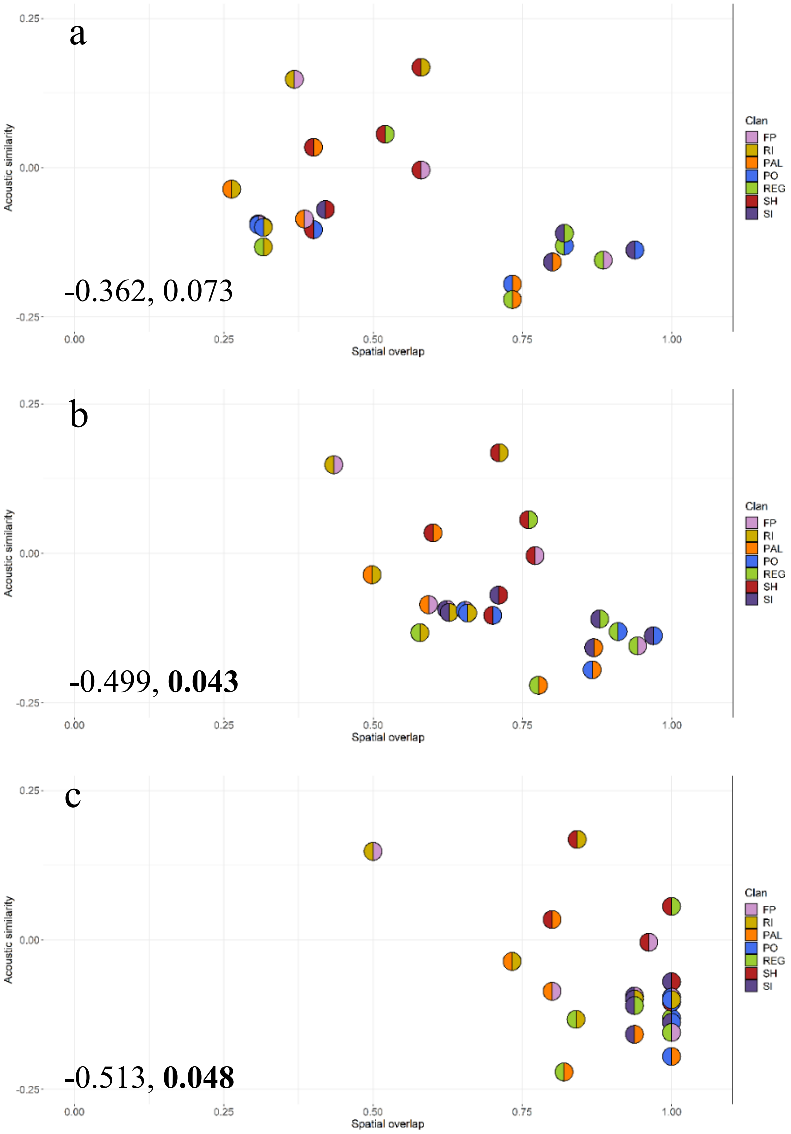


*Figure 4.5* – Clan spatial overlap vs. $acsim_{IDbt}$. Each circle represents a pair of clans (see Figure 4.1 for clan abbreviations). For each panel—(a) minimum spatial overlap; (b) mean spatial overlap; and (c) maximum spatial overlap—the Mantel test matrix correlation and p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that $acsim_{IDbt}$ decreases as clan spatial overlap increases.

In contrast, plots of correlations in usage of non-identity codas between pairs of clans ($acsim_{nonIDbt}$) against the spatial overlap of the clans showed no obvious trend, and none of the Mantel test p-values approached significance (Figure 4.6). This suggests that, in contrast to identity codas, similarity in non-identity coda usage between clans does not systematically change as their spatial overlap increases. The magnitude of the Mantel test matrix correlation values was always larger for $acsim_{IDbt}$ (Figure 4.5) compared to $acsim_{nonIDbt}$ (Figure 4.6), which likely explains why the $acsim_{allbt}$ plots show a decreasing trend with increasing spatial overlap like the $acsim_{IDbt}$ plots, but it is not significant (Figure S4.11). The Mantel test results of clan spatial overlap against $acsim_{diffbt}$ were not significant (although maximum overlap vs. $acsim_{diffbt}$ was marginally significant with p=0.062; Figure S4.12).

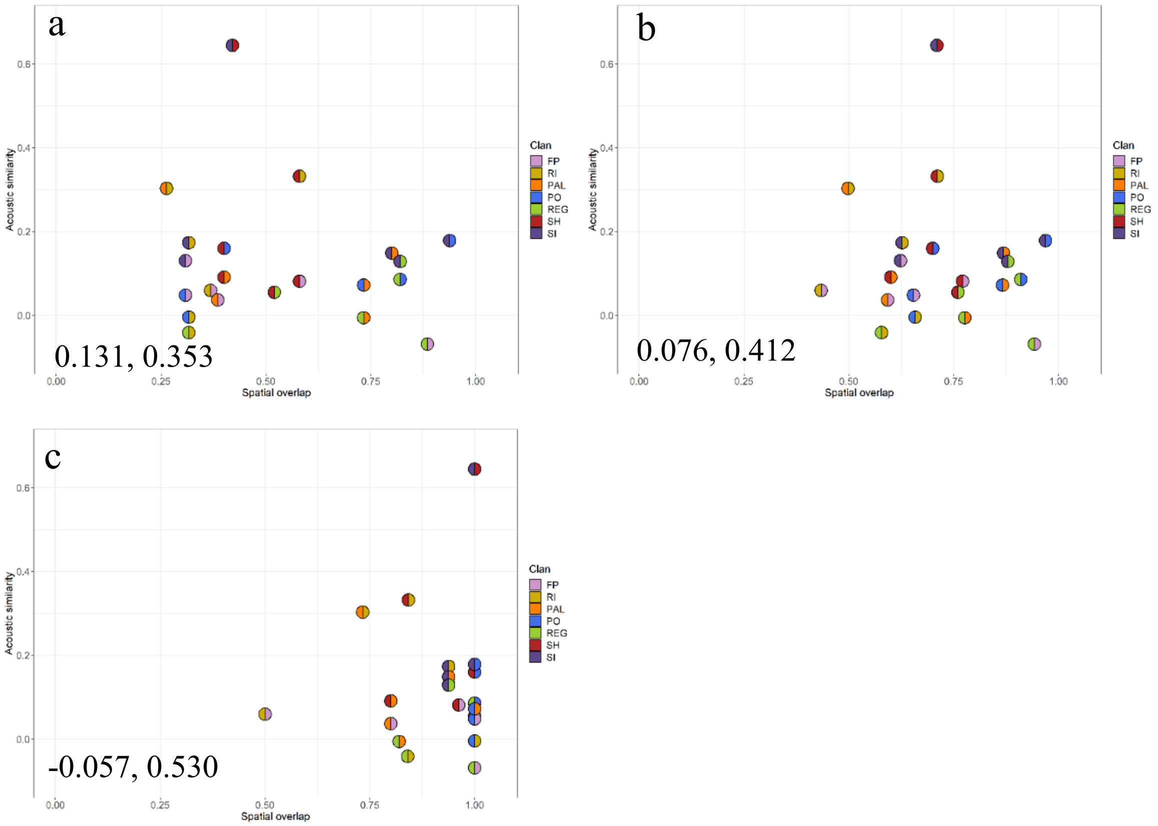


*Figure 4.6* – Clan spatial overlap vs. $acsim_{nonIDbt}$. Each circle represents a pair of clans (see Figure 4.1 for clan abbreviations). For each panel—(a) minimum spatial overlap; (b) mean spatial overlap; and (c) maximum spatial overlap—the Mantel test matrix correlation and p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that $acsim_{nonIDbt}$ decreases as clan spatial overlap increases.

Thus, we could not reject the null hypothesis that acoustic similarity between all clans, whether calculated using just identity codas or just non-identity codas, drops similarly with increasing clan spatial overlap across all clans, but the results from $acsim_{IDbt}$ and $acsim_{nonIDbt}$ suggest that identity coda usage is modulated as clan spatial overlap increases, while non-identity coda usage is not.

Results using an alternative way of calculating between-clan acoustic similarity (by averaging the correlation values for every pairwise combination of repertoires in two clans; $acsim_{btAlt}$) showed similar trends, although p-values changed (Figures S4.13–S4.16). The multidimensional scaling plots illustrate that between-clan acoustic similarity can differ depending on which coda types are examines (Figures S4.17/S4.18).

## 4.5 – Discussion

By leveraging regional research efforts, we compiled detailed repertoires of sperm whale codas from across the Pacific. At least seven sperm whale cultural clans likely inhabit the Pacific Ocean, and these clans: have variable and overlapping distributions, are generally well distinguished by identity codas, and vocalize and behave differently, with some clan vocal characteristics modulated by clan spatial overlap.

### *4.5.1 – Distribution of clans in the Pacific Ocean*

Each of the seven clans has a different distribution, with order of magnitude differences in some clan spatial spans (e.g. 1,000 km for Plus-One vs. 10,000 km for Short). Many regions have sympatric clans (Figure 4.1). The general decrease in the number of detected clans in regions as the number of repertoires/codas decreases suggests that sympatric clans may be the rule rather than the exception across the Pacific, and that there may be additional clans in un(der)sampled regions (Table S4.6).

Two features of Pacific clan distributions are especially fascinating: that many clans span huge swaths of the ocean, and that all clans spatially overlap with at least one other clan in some portion of their detected range. This first feature is also seen in blue whales (*Balaenoptera musculus*), which are divided into populations based on cultural songs (McDonald et al., 2006, 2009). However, while some blue whale populations within an ocean basin spatially overlap with others, this is not the case for all, and the distribution of blue whale populations in the Pacific Ocean is more modular than that of sperm whale clans (McDonald et al., 2006). Orca (*Orcinus orca*) ecotypes have culturally-inherited behaviors (including vocal dialects), live in sympatry in several parts of the world, and can have very large ranges (de Bruyn et al., 2013; Riesch et al., 2012), but some ecotypes are also well on their way to being different species (LeDuc et al., 2008), while sperm whales are a single species with little distinction in nuclear DNA between clans (Konrad et al., 2018). Indeed, the distribution patterns observed in Pacific sperm whale clans most closely echo some of the patterns seen in human ethnolinguistic diversity, including uneven distribution (i.e. high diversity in some areas and low diversity in others; some groups having very much larger ranges than others), a latitudinal gradient (i.e. greater diversity near the equator than the poles), and frequent regions of overlap (i.e. multiple ethnolinguistic groups/clans in the same area) (Currie & Mace, 2012). The global distribution of human ethnolinguistic groups has also been driven by factors like political complexity, environmental productivity and heterogeneity, and subsistence strategies (Currie & Mace, 2009, 2012; Honkola et al., 2018), some of which have homologues or proxies in sperm whales (e.g. social complexity, ocean productivity and heterogeneity, foraging strategy) that could be investigated in future work.

### *4.5.2 – Clan identity codas*

All repertoires recorded across the Pacific could be assigned to clans based on identity codas made frequently by the clan and rarely outside it. Within clans, we saw a trend of decreasing acoustic similarity with increasing distance between repertoire recording locations. An exception was the most geographically restricted clan, the Plus-One clan, which spanned ~1,000 km (the approximate annual home range of eastern tropical Pacific sperm whales; Whitehead, 2001; Whitehead et al., 2008). Isolation by distance has been documented in culturally transmitted attributes of other species, such as human material culture (Lycett, 2019) and red-faced cisticola (*Cisticola erythrops*) song (Benedict & Bowie, 2009). In sperm whales, this decrease seems to be driven more by decreases in usage similarity of non-identity codas compared to identity codas. Our analyses thus suggest that different coda types experience different evolutionary pressures, with more stability over geographic space in identity codas and more drift in non-identity codas. Variable selection pressure depending on call type, function, or frequency has been observed in communication systems of other taxa as well, including humans (Pagel et al., 2007), orcas (Filatova, Deecke, et al., 2012), and passerine birds (e.g. Daurian redstarts, *Phoenicurus auroreus,* Lee et al., 2019; Savannah sparrows, *Passerculus sandwichiensis,* Williams et al., 2013).

While we cannot know for certain how whales perceive codas or label the identity of conspecifics, our results provide the first quantitative support for identity codas having a clan identification function, given that their usage reflects patterns apparent in human ethnic group markers (Boyd & Richerson, 1987; McElreath et al., 2003). These patterns likely arise from the selection for identity signals that enable social assortment and cooperation in sperm whales, as has been shown for humans (Boyd & Richerson, 1987; Moffett, 2013) and suggested for other taxa, including chimpanzees (*Pan troglodytes*) (Crockford et al., 2004), short-finned pilot whales (*Globicephala macrorhynchus*) (Van Cise et al., 2018), and orcas (Filatova, Deecke, et al., 2012). Further confirmation could be provided by playback experiments, where whales are exposed to same-clan, different-sympatric clan, or different-allopatric clan identity codas and their behavioral response is measured, but such studies are logistically challenging for wild marine mammals (Deecke, 2006).

### *4.5.3 – Different clans vocalize and behave differently*

The emergent picture from our analyses is that different clans do things differently. Pacific clans vary, sometimes drastically, in metrics like number of identity codas, coda click length preferences, within-clan correlation of repertoires, and geographic range. Some commonalities also emerge, such as rhythmic motifs in the identity codas of most clans (although exact motifs vary by clan). A more in-depth characterization of the full coda repertoires of clans will help determine how far these motifs extend beyond identity codas.

The vocal and distributional differences we documented are likely just the tip of the iceberg when it comes to phenotypic differences between clans in the Pacific. From research off the Galápagos Islands and in the eastern Caribbean, we know that phenotypic differences among clans extend beyond vocal dialects and can have fitness consequences (Marcoux, Rendell, et al., 2007; Marcoux, Whitehead, et al., 2007; Vachon et al., submitted; Whitehead & Rendell, 2004). For example, Plus-One and Regular clan whales have different foraging strategies that resulted in differential foraging success during El Niño years and 'normal' years (Whitehead & Rendell, 2004). Interestingly, the Plus-One clan had significantly greater displacements over 6-hour to 5-year scales than the more widely distributed Regular clan (Whitehead et al., 2008), which suggests that the major differences in clan distribution seen across the Pacific here may not be predicted by ranging behavior (otherwise the opposite trend would be expected). These differences may instead relate more to *perceived* barriers to movement, which could include cultural barriers, like conformity.

Our choice to restrict analyses to 3–10-click codas could potentially obscure some clan boundaries. For example, two-click codas are made in recordings from the Mariana Islands and the Ogasawara Islands of Japan but were not analyzed because they have been inconsistently marked across datasets; these very short codas may be culturally relevant, and their omission could explain why some of the Mariana Islands repertoires have low within-clan correlations (Figure 4.1). Codas with three or four clicks have low dimensionality (and potentially lower information content, as seen in human words; Piantadosi et al., 2011) and form more diffuse clusters than codas with more clicks, which may have contributed to the uncertainty of clan divisions for repertoires dominated by codas with fewer clicks

### *4.5.4 – Modulation of clan vocal behavior by clan overlap*

Acoustic similarity based on identity coda usage decreased as clan spatial overlap increased, whereas no trend was seen in non-identity coda usage. This supports the hypothesis that identity codas are used as symbolic markers of clan identity, and aligns with simulation work by McElreath et al. (2003) that found that group/ethnic differences are strongest at boundaries for symbolically marked groups. In the case of Pacific sperm whales, more spatially overlapped clans appear to be "more marked" than less spatially overlapped clans which could imply that between-clan interactions have increased selection on usage of identity markers (in this case, identity codas) (McElreath et al., 2003). This fits with cross-species research showing that the dialects of groups or species in sympatry are more distinct than in allopatry (e.g. southern wood crickets, *Gryllus fultoni*, Jang & Gerhardt, 2006; flycatchers, *Ficedula* spp., Haavie et al., 2004; green tree frogs, *Hyla* spp., Höbel & Gerhardt, 2003; chimpanzees, Crockford et al., 2004) due to heightened selection for divergence in sympatry (often to avoid mismating) (Pfennig & Pfennig, 2009).

No single coda type is unique to a single clan, which means that clans occasionally make each other's identity codas. The between-clans analysis results suggest that clans make other clans' identity codas *less* when in areas of higher spatial overlap. This adjustment would theoretically improve their ability to quickly discern if nearby whales are in their clan or a different clan, which could guide decisions about social assortment or avoidance (Cantor & Whitehead, 2013, 2015).

### *4.5.5 – Future directions*

There are many avenues for future work that can build on our collation and analysis of a very large, and large-scale, coda dataset. Deployment of autonomous recorders in areas that are difficult to access but frequented by sperm whales will improve our coverage of the Pacific Ocean. Additional work in regions like Palau and the Sea of Cortez will be clarifying, since the single repertoires recorded in those locations have low within-clan correlations; this may be a sampling artifact or could hint at additional, uncharacterized clans in these regions. Comparisons with results from other oceans will be interesting as well. Sperm whales in the eastern Caribbean and Mediterranean show substantial and robust differences in social structure, ranging behavior, and vocalizations compared to those in the eastern tropical Pacific (Vachon et al., submitted; Whitehead et al., 2012; Rendell & Frantzis, 2016). These extreme differences suggest that different selection pressures may be operating in different bodies of water, which could lead to within-clan and between-clan acoustic trends that differ from those documented in the Pacific.

When quantifying selection pressure, understanding the characteristics of the basal or source population can be very informative, and is required to draw conclusions about if and how between-group interactions have impacted selection on group identity markers (McElreath et al., 2003). For example, phylogenetic models informed by cognate evolution rates can be used to infer ancestral dialects in human linguistics, which can help track historic migration patterns of speakers (Bouckaert et al., 2012; Currie et al., 2013). Results from our analyses provide preliminary support for the Short clan dialect as basal in the Pacific Ocean (Discussion S4.2 in Appendix C), but a detailed understanding of how codas evolve over time is necessary before we can determine the 'ancestral' sperm whale dialect.

In tandem with geographically expanded research efforts, we need to develop automated methods for extracting codas from recordings, as our ability to collect vast quantities of acoustic data has outpaced our ability to analyze it. This is a key objective of the newly launched Cetacean Translation Initiative (Andreas et al., 2021).

### *4.5.6 – Implications*

Symbolic marking has rarely been documented outside of humans, but we found evidence that identity codas act as symbolic markers of sperm whale clan identity in the Pacific Ocean. This discovery provides insight on how different clans maintain distinctiveness and boundaries in areas of sympatry, and further highlights similarities between human and sperm whale culture.

Despite culture clearly structuring sperm whale populations, the species is managed using broad geographic stocks that lump multiple sympatric clans together. With the conservation implications of animal culture increasingly being acknowledged (Brakes et al., 2019, 2021; Keith & Bull, 2017), the Convention on Migratory Species is developing a concerted action plan for the well-studied eastern tropical Pacific sperm whales to determine whether different clans should be managed as separate, socially significant units (Convention on Migratory Species, 2017). Relevant to this action plan, the present research updates the number of known eastern tropical Pacific sperm whale clans (from four to seven) and the Convention on Migratory Species Range States in which different clans have been identified (Table S4.8). This illustrates how results of analyses like ours can be directly applied to conservation policy and potentially aid managers. Importantly, the geographic scope of our analysis extends well beyond the eastern tropical Pacific, providing, for the first time, benchmarks of sperm whale cultural diversity across the Pacific Ocean.

## 4.6 – Conclusions

We showed wide-scale cultural structuring of sperm whale populations across the largest ocean on Earth. The distributional scale and trends resemble patterns found in human ethnic groups. Our results suggest that sperm whales in the Pacific use identity codas as symbolic markers of cultural clan identity and alter their vocal behavior in ways that reaffirm this identity in areas of high spatial overlap. This first ‘census’ of Pacific sperm whale culture provides baseline information that can facilitate culturally-sensitive conservation and management of this deep-sea keystone species.

## 4.7 – Acknowledgements

This chapter emanates from the Global Coda Dialect Project, a consortium of scientists conducting sperm whale acoustics research worldwide. We are grateful to: the individuals that helped collect acoustic data and the organizations that funded regional research efforts; Noor Youssef, Scott McCain, and Elizabeth Zwamborn for conversations that helped orient this work; Kristian Beedholm for creating and sharing Coda Sorter; Robin Baird and Katie Kowarski for connecting us with individuals who had sperm whale data; and Héloïse Frouin-Mouy, Lynn Lee, Jasper Kanes, Ocean Networks Canada, and Richard Dewey for sharing recordings that were ultimately not included in this publication. Our coauthor and colleague, Thomas Norris, passed away on September 9, 2020. We are grateful to have had the chance to work on this project with him.

# CHAPTER 5

# STABILITY AND EVOLUTION IN SPERM WHALE CULTURAL DIALECTS[22]

> "Listen. To live is to be marked. To live is to change, to acquire the words of a story, and that is the only celebration we mortals really know."
>
> —Barbara Kingsolver

## 5.1 – Abstract

Understanding the stability of population-wide vocalizations over time can help illuminate the underlying function of calls, as well as transmission biases and directions of information flow in communication systems. For species that exhibit culture, such as sperm whales, cultural drift and cultural selection can cause vocalizations to change within the lifespan of individuals. In social situations, sperm whales communicate using patterned series of clicks, called codas. Different cultural groups of whales, called vocal clans, prefer different coda types. We quantified the stability of fine-scale structural coda parameters (total duration, inter-click interval ratios) over 11-to-40-year timespans for five well-sampled vocal clans in three regions: the Regular and Plus-One clans off the Galápagos Islands; the Eastern Caribbean 1 and Eastern Caribbean 2 clans off the Lesser Antilles; and the Mediterranean clan off the Balearic Islands. Within-type coda similarity generally decreased over time across clans, but exact levels of temporal stability differed by coda type. For coda types that had significantly changed, the magnitude and direction of change varied, with some codas getting slower over time and others getting faster. Collectively, these results emphasize that the sperm whale social communication system is dynamic and suggest that codas are susceptible to cultural drift, although additional

---

[22] This chapter is in collaboration with: Shane Gero, Luke Rendell, Maurício Cantor, Félicia Vachon, and Hal Whitehead. TAH analyzed the data, interpreted the results, and drafted the chapter, with input from SG, LR, and HW. All collaborators collected sperm whale coda data.

research accounting for behavioral context during coda production and quantifying sperm whale auditory discrimination abilities is necessary. Nevertheless, the changes we documented were extremely fine-scale, and suggest that coda dialects can be used to facilitate conservation of sperm whales at the vocal clan level.

## 5.2 – Introduction

There are well-established links between social and vocal complexity for many taxa (e.g. non-human primates, Bouchet et al., 2013; birds, Freeberg, 2006; bats, Wilkinson, 2003). To fully grasp the communicative function of different vocalizations in social species, however, we need to quantify the stability of those vocalizations over time. This type of research is foundational to understanding the role vocal communication plays in a species' life history and which genetic, environmental, social, and/or cultural factors shape the evolution of vocalizations. From a conservation perspective, passive acoustic monitoring can be a valuable tool for identifying and tracking different populations of animals, but only if they can be reliably distinguished by temporally stable acoustic signatures (Hersh et al., 2021; Weirathmueller et al., 2017).

For some taxa, vocalizations are extremely stable over time. For example, Darwin's finches (*Geospiza fortis*) on Santa Cruz Island showed stability in various song type parameters (including number of notes, song duration, and note frequency) over 40 years (Goodale & Podos, 2010). The diversity of call types used in harp seal (*Pagophilus groenlandicus*) underwater vocal repertoires has not changed for decades (Serrano & Terhune, 2002; Terhune, 1994), and bottlenose dolphin (*Tursiops truncatus*) signature whistles, which convey individual identity, are stable for at least 12 years (Sayigh et al., 1990). At the other end of the continuum are species whose vocalizations are extremely labile over time. Humpback whale (*Megaptera novaeangliae*) songs undergo 'cultural revolutions', much like human music (Mauch et al., 2015), with population-wide adoption of new songs in as little as two years (Garland et al., 2011; Garland & McGregor, 2020). In the northeast Pacific Ocean, fin whales (*Balaenoptera physalus*) gradually shifted from singing 'singlet' songs (with isochronously spaced calls) to singing 'doublet' songs (with heterochronously spaced calls) over a decade (Weirathmueller et al., 2017). Even within a species, vocalization stability over time can vary. For example, hermit warblers (*Septophaga occidentalis*) in northern Oregon exhibited significant changes in type I songs over an 11-year period (including syllable loss, phrase splitting, and phrase lengthening), while the songs of warblers in southern Oregon did not change (Janes & Ryker, 2013). Fine scale structural changes have also been documented in some, but not all, northern and southern resident orca (*Orcinus orca*) discrete calls (Deecke et al., 2000; Wieland et al., 2010). Why do some stereotyped vocalizations change over time within a population, while others do not?

Whether a vocalization changes over time or not can depend on its function, the selective forces acting on it, and the time period being considered (Rekdahl et al., 2013). Several researchers have hypothesized that vocalizations that play important roles in social interactions, such as facilitating group contact, cohesion, or coordination, are likely to remain stable over time (Rekdahl et al., 2013; Riesch et al., 2006). On annual timescales, population-level changes in vocalizations may relate to reproductive hormone levels, as has been suggested for fin whales (Oleson et al., 2014). Population vocalizations can also change over time as an adaptation to changing levels of environmental noise. This has been documented for various taxa in response to increases in anthropogenic noise (e.g. great tits, *Parus major,* Slabbekoorn & Peet, 2003; bananaquits, *Coereba flaveola*, Winandy et al., 2021; right whales, *Eubalaena* spp., Parks et al., 2007; orcas, Foote et al., 2004) and can manifest as changes in call duration and/or frequency. For species that exhibit culture—information or behavior that is shared within a community and acquired from conspecifics through social learning (Whitehead and Rendell, 2014)—cultural drift and cultural selection can also cause vocalizations to change. In such cases, different transmission biases (e.g. homophily, conformism, symbolic marking) and directions of information flow (e.g. horizontal, vertical, oblique) can impact the scale and rates of change in vocalizations, leaving signatures of evolutionary processes/drivers behind (Cantor et al., 2015). For example, the cultural revolutions observed in humpback whale song result from horizontal transmission of new song variants, with extreme cultural conformity coupled with a clear "quest for novelty" (Garland et al., 2011; Garland & McGregor, 2020; Rekdahl et al., 2013). McDonald et al. (2009) hypothesized that the worldwide decline in blue whale (*Balaenoptera musculus*) song frequency was at least partially facilitated by strong cultural conformity, while the changes observed in orca discrete calls were most parsimoniously explained by cultural drift with horizontal transmission of call modifications (Deecke et al., 2000).

One species for which culture is a pervasive aspect of life is the sperm whale (*Physeter macrocephalus*). In social situations, sperm whales communicate using patterned series of clicks, called codas. Different types of codas are produced by varying the number of clicks and the inter-click intervals (ICIs), and different social units of whales (i.e. matrilineal groups of females and juveniles) have only been observed interacting if they use similar types of codas (i.e. have similar dialects) (Rendell & Whitehead, 2003b). These biased social unit interactions give rise to a culturally determined level of population structure: the vocal clan. Recent work found evidence that certain 'identity codas' are highly characteristic of particular vocal clans (Hersh et al., 2021) and may act as symbolic markers of clan identity (chapter 4), while other coda types ('non-identity codas') are made by multiple clans.

Previous work on eastern tropical Pacific (Rendell & Whitehead, 2005) and eastern Caribbean (Gero, Whitehead, et al., 2016) sperm whale social units found no evidence of change in coda output over timespans ranging from one day to six years, but such spans are a small fraction of an adult female sperm whale's ~60 year lifespan (Rice, 1989) and may not be appropriate for detecting change. Rendell & Whitehead (2005) also standardized codas by total duration prior to analysis, which preserved coda rhythm but not tempo, but more recent research has shown that tempo is an important coda feature (Antunes et al., 2011; Frantzis & Alexiadou, 2008; Gero, Whitehead, et al., 2016).

With additional years' worth of acoustic data now available, coda type stability over time is worth revisiting for several reasons. Modelling work by Cantor et al. (2015) showed that sperm whale vocal clans likely originated from cultural transmission via biased social learning of codas, rather than from individual learning, genetic drift, or cultural drift. Homophily and conformism were important transmission biases, but symbolic marking was not (Cantor et al., 2015). However, the transmission processes and biases that were key to clan origination may differ from those that facilitate clan maintenance, particularly in areas with sympatric clans. Cultural drift has been implicated in vocalization evolution and acoustic clan maintenance in orcas, which share many behavioral and life history similarities with sperm whales (e.g. matrilineal social structure, cultural xenophobia, vocal culture) (Deecke et al., 2000; Riesch et al., 2006; Yurk et al., 2002). Coupled with recent advances in our understanding of identity codas, non-identity codas, and symbolic marking in sperm whales (chapter 4; Hersh et al. 2021), we can now take a more nuanced approach to the question of dialect stability over time. Additionally, researchers have advocated for using vocal clans as evolutionarily significant units (ESUs) in sperm whale conservation since clans were first described in 2003 (Brakes et al., 2019; Rendell & Whitehead, 2003b). To do so, it is imperative that we first understand if and how the building blocks of clan dialects—identity and non-identity codas—change over time (Weirathmueller et al., 2017).

Here, we quantify within-coda type stability over time for different sperm whale vocal clans. Recent analyses have shown a within-clan trend of decreasing acoustic similarity with increasing distance (>1,000 km[23]) between repertoire recording locations (chapter 4). To account for this potential spatial drift, we conducted analyses of temporal stability on codas from well-sampled clans in three study regions that each spanned <1,000 km: the Galápagos Islands (Plus-One and Regular clans), the Lesser Antilles (Eastern Caribbean (EC) 1 and EC2 clans), and the Balearic Islands (Mediterranean clan). In contrast to past work, our approach focuses on fine-scale structural features of coda types themselves (i.e. duration, ICI ratios) rather than broad-scale trends in coda output. The total timespan of recordings included here ranges from 11 to 40 years depending on the clan, representing significant increases (83–2,800%) compared to past work (Gero, Whitehead, et al., 2016; Rendell & Whitehead, 2005), and arguably more meaningful timescales in the life of an adult female sperm whale.

We hypothesize that identity coda types have not changed over time, given that such codas may be symbolic markers of clan identity (chapter 4). In contrast, we hypothesize that some non-identity coda types have changed over time due to nondirectional cultural drift, since non-identity codas may be under less selective pressure for stability because they are not the primary indicators of cultural group identity (chapter 4).

---

[23] This is the approximate annual home range of tropical Pacific sperm whales (Whitehead et al., 2008).

## 5.3 – Methods

### *5.3.1 – Data collection and coda extraction*

Protocols for collecting and extracting codas from sperm whale acoustic recordings have been extensively detailed in past work[24]. Briefly, various hydrophones and recording systems were used to record sperm whales (primarily adult females and juveniles) in the eastern tropical Pacific Ocean, Caribbean Sea, and Mediterranean Sea between 1978 and 2020. Recordings were manually audited for the presence of codas, and descriptive coda metrics were extracted (including ICIs measured in s and number of clicks), typically using 'Rainbow Click' (Gillespie, 1997) or 'Coda Sorter' (a custom software implemented in LabView and run in MATLAB version R 2020a). While differences in recording protocols/equipment can significantly affect spectral features of vocalizations (Weirathmueller et al., 2017), temporal features like those we examined here (i.e. ICIs, total duration) are more robust to such variation. All ICIs were rounded to three decimal places to match the maximum time resolution of our analysis equipment (0.001 s; Rendell & Whitehead 2003b). Extracted codas with 3–10 clicks were divided into repertoires by recording day within regions.

### *5.3.2 – Coda typing and clan assignment*

The 'IDcall' program (Hersh et al., 2021) was used to classify codas into types and assign repertoires to clans. This was done separately for codas recorded in Pacific vs. Atlantic/Mediterranean waters. In IDcall, codas with the same number of clicks are classified into types using parsimonious mixtures of multivariate contaminated normal distributions. The types themselves are continuous (i.e. not discrete), and each coda is assigned a probability of belonging to each type. Outliers are automatically identified (see Punzo et al., 2018 for details). Next, repertoires of codas are hierarchically clustered based on similarities in coda type usage, such that repertoires comprised of similar coda types and with similar usage proportions cluster together in the resultant dendrogram. If a cluster of repertoires is united by high usage of an 'identity coda' (i.e. a coda type made

---

[24] For the eastern tropical Pacific, see Cantor et al. (2016), Hersh et al. (2021), Rendell & Whitehead (2003b), and Weilgart & Whitehead (1997); for the Caribbean Sea, see Vachon et al. (submitted), Watkins & Moore (1982), Gero, Bøttcher, et al. (2016), Gero, Whitehead, et al. (2016), Tønnesen et al. (2018), and Hersh et al. (2021); for the Mediterranean Sea, see Pirotta et al. (2011) and Rendell et al. (2014).

frequently in those repertoires but rarely in any others), that cluster will be designated an 'identity clade' (i.e. putative clan).

As in chapter 4, we fitted 2:15 mixture components to the coda data and used a subset of the 14 possible models (determined using the 'mclust' R package, Scrucca et al. 2016; see Punzo & McNicholas, 2016 for model descriptions) during coda classification. All codas were included in the classification step, but only repertoires with at least 25 codas were assigned to clans during hierarchical clustering (Rendell & Whitehead, 2005). During hierarchical clustering, we set both IDcall parameters—*minrep* (i.e. the minimum number of repertoires required for an identity clade to form) and *critfact* (i.e. the factor by which a coda type must be made more in a given clade compared to all other clades for it to become an identity coda)—to 10 (Hersh et al., 2021). These conservative parameter settings prevent the formation of putative clans comprised of just a few repertoires and help ensure that identity codas are highly unique to the putative clan that produces them.

### *5.3.3 – Dataset restrictions*

Coda-to-type classification certainty and repertoire-to-clan assignment certainty were evaluated using posterior probabilities of codas belonging to types (calculated in IDcall using the 'ContaminatedMixt' R Package; Punzo et al., 2018) and repertoires belonging to clans (calculated via the IDcall 'postprob' extension and using all coda types; Whitehead & Hersh, submitted), respectively. Codas were omitted if they were designated as outliers during classification; had a type classification posterior probability <0.9; or belonged to a repertoire that had a clan assignment posterior probability <0.9. Of the remaining codas, only well-sampled types (i.e. those with at least 25 codas in at least two different years) were included in within-type analyses of temporal stability. These steps ensured that misclassification of codas to types or misassignment of repertoires to clans did not lead to erroneous conclusions about temporal stability in codas. All analyses were carried out within types within clans.

### *5.3.4 – Within-type analyses of temporal stability*

To quantify temporal stability, we first calculated the Euclidean distance between the ICIs of every pair of codas of a given type. To account for the fact that codas from the same repertoire (i.e. recording day) may not be independent (Cantor et al., 2016), we replaced the Euclidean distance of same-repertoire codas with NaNs. We then plotted the Euclidean distances (excluding NaNs) against the number of elapsed years between coda recordings and fit a linear regression. A positive slope suggests that codas have become more dissimilar over time, while a negative slope suggests that codas have become more similar over time. Using the 'cultevo' R package (Stadler, 2018), we ran one-tailed Mantel tests with 10,000 permutations to assess the significance ($p \leq 0.05$) of any positive slopes. To better understand whether certain coda type analyses were underpowered, we plotted the coda sample size of the second largest year against slope and inspected the plot for trends. We also ran a Welch's two sample t-test (using the 'stats' R package) to compare the Mantel test matrix correlation coefficients of identity and non-identity codas.

Those analyses address *whether* codas within types have changed over time (i.e. become more or less dispersed in multivariate ICI space), but not *how* they have changed. For coda types with significant Mantel test results (i.e. significant evidence of change over time), we used linear mixed-effects models to see whether coda type total duration (i.e. the sum of the ICIs) changed over time. For each type, the model included year as a quantitative fixed effect and repertoire as a random effect (once again addressing potential nonindependence of codas from the same repertoire; Cantor et al., 2016). If the random effect was not significant, we removed it and ran a linear model instead[25]. We also ran principal component analysis (PCA) on the ICIs for each coda type and used the first two principal components, PC1 and PC2, as response variables in additional models. To determine which coda feature(s) PC1 and PC2 represented, we inspected the PCA loadings. To examine the extent to which years with small sample sizes contributed to the results, we also ran the clan analyses using *only* years with at least 25 codas of a given type. In total, there were thus two forms of analysis for each coda type: at the clan level using codas from all years ($\text{clan}_{\text{allyears}}$) and at the clan level using codas from well-

---

[25] For brevity, we refer to the linear mixed-effects models and the linear models as 'models' from this point onward but see the referenced Tables for additional information on which type of model was run.

sampled years (clan$_{\text{largeyears}}$). All models were run in MATLAB (version R 2020A) using the ‘fitlme’ function.

## 5.4 – Results

### *5.4.1 – Coda dataset*

In total, 39,263 codas recorded between 1978 and 2020 (comprising 325 repertoires) were input to IDcall (Table S5.1). This dataset includes the regions and codas analyzed in (Hersh et al., 2021) but differs in a few key ways: codas with 3–10 clicks are included (compared to codas with 3–9 clicks in Hersh et al., 2021); additional codas from the Lesser Antilles have been incorporated (including over 6,000 codas recorded in 2019 and 2020; Vachon et al., submitted); codas (n=217) from various years were added to the Balearic Islands dataset; and codas were divided into repertoires by recording day within regions, whereas Hersh et al. (2021) used recording year within regions.

### *5.4.2 – Clans and identity codas*

Preliminary data exploration with the ‘mclust’ R package (Scrucca et al., 2016) showed that four of the 14 possible models consistently best fit the coda data across click lengths, as determined by the Bayesian Information Criterion (Figures S5.1/S5.2). These four models—VEE, VVE, VEV, and VVV (see Punzo & McNicholas (2016) for model descriptions)—were also the best fitting models in chapter 4 and were included as options during the call classification stage of IDcall.

IDcall detected 91 coda types (18 identity) in the Atlantic/Mediterranean dataset and 104 coda types (28 identity) in the Pacific dataset. Setting both *critfact* and *minrep* to 10 resulted in an Atlantic/Mediterranean dendrogram (cophenetic correlation coefficient=0.877) with the same three clans found in Hersh et al. (2021) (Figures S5.3/S5.4, Table S5.2). The six clans in the Pacific dendrogram (cophenetic correlation coefficient=0.930) match those found in past work (Rendell & Whitehead 2003b; Cantor et al., 2016; Hersh et al., 2021; chapter 4) (Figures S5.5/S5.6, Table S5.2). In both dendrograms, the clan identity codas also mirror those identified in previous studies (Cantor et al., 2016; Gero, Bøttcher, et al., 2016; Pavan et al., 2000; Rendell & Whitehead, 2003b; chapter 4).

For the five clans of interest—EC1, EC2, Mediterranean, Plus-One, and Regular—the total timespan of recordings ranged from 11 to 40 years. The various dataset restriction steps (e.g. removing codas designated as outliers and codas/repertoires with low posterior probabilities) collectively reduced the clan coda sample sizes by 53.4–77.2% (Table S5.3).

### *5.4.3 – Within-type analyses of temporal stability*

#### *5.4.3.1 – Mantel test results for Euclidean distance measure*

Across clans, we investigated stability over time in 18 coda types (Table 5.1, Figures S5.7–5.11). Slopes and matrix correlation coefficients were positive for 14 types (i.e. the Euclidean distance between codas increased over time), meaning that a majority (77.8%) trended towards becoming less similar over time (Table 5.1). The null hypothesis of no significant change over time was rejected for six identity codas and two non-identity codas from four of the five clans (Table 5.1), and these coda types underwent additional analyses (i.e. modelling). Neither of the EC2 clans' codas had significant Mantel test results, although both had positive slopes (Table 5.1). Across all clans, the matrix correlation coefficients for identity and non-identity codas did not significantly differ (Welch's two sample t-test, $p=0.704$).

Plotting the coda sample size of the second largest year against the slope of the Euclidean distance plot shows some interesting patterns (Figure S5.12). For the three Regular clan coda types (64, 73, 810), the sample size of the second largest year is <50 codas, but the effect size is quite large and all three Mantel test results were significant. This suggests that those coda types truly have changed in some way over time. Other coda types, like 51 and 511 from the EC1 clan and 514 from the EC2 clan, have large second year sample sizes (>100 codas). These robust sample sizes lend confidence to the Mantel test results, which suggest that types 511 and 514 have not significantly changed over time but type 51 has. For many of the coda types, however, both the second largest year sample size and the effect size are small; it is difficult to conclusively say if these types truly did or did not change over time, or if these analyses were underpowered.

*Table 5.1.* Coda types included in within-type analyses of temporal stability. Clan abbreviations are: EC1=Eastern Caribbean 1, EC2=Eastern Caribbean 2, MED=Mediterranean, PO=Plus-One, and REG=Regular. The ‘Coda type’ column gives the numeric coda type codes; for type names, see Table S5.4 (EC1, EC2, MED) and Table S5.5 (PO, REG). Note that numeric coda type codes may overlap between the two trees (e.g. type 515) but only indicate the same coda type *within* a tree. Stars in the ‘Coda type’ column denote identity coda types. ‘Recording year span’ includes the first and last years with codas, and the number in parentheses is the recording year span if only years with ≥25 codas are included (i.e. the bolded years in the ‘Recording years’ column). The line slope was calculated from the Euclidean distance plot. Mantel test p-value notation is as follows: 0=***, 0.001=**, 0.01=*. All values >0.01 are typed out, with significant p-values (≤0.05) bolded. Gray shading denotes regions included in the Pacific tree; no shading denotes regions included in the Atlantic/Mediterranean tree.

| **Clan** | **Coda type** | **Recording year span** | **Recording years** | **Total codas** | **Line slope (s/year)** | **Mantel test matrix correlation coefficient** | **Mantel test p-value** |
|---|---|---|---|---|---|---|---|
| EC1 | 46 | 15 (10) | 20: 05, 08, 09, **10**, **19** | 191 | 0.00041 | 0.222 | *** |
| | 51* | 40 (40) | 19: **81**, 83, **84**, 87, **90**, 94<br>20: **05**, 07, **08**, 09, **10**, 11, **14**, **15**, **16**, **18**, **19**, **20** | 1,972 | 0.00025 | 0.063 | *** |
| | 52 | 13 (3) | 20: **08**, 09, **10**, 11, 14, 15, 16, 17, 18, 19, 20 | 140 | -0.00007 | -0.021 | 0.760 |
| | 54* | 40 (13) | 19: 81, 90<br>20: 07, **08**, 09, **10**, 11, **15**, 16, 17, 18, 19, **20** | 215 | 0.00012 | 0.077 | **0.020** |
| | 55 | 40 (26) | 19: 81, 83, 84, 87, **90**, 94<br>20: 05, 08, 09, **10**, 11, 14, **15**, 18, 19, 20 | 344 | -0.00002 | -0.020 | 0.887 |
| | 57* | 40 (31) | 19: 81, 84, 87, **90**, 94<br>20: 05, 07, **08**, 09, **10**, 11, 14, **15**, 16, **18**, 19, **20** | 482 | -0.00007 | -0.032 | 0.932 |
| | 511* | 27 (13) | 19: 94<br>20: 05, 07, **08**, 09, **10**, 11, **14**, **15**, **16**, **17**, 18, **19**, 20 | 879 | 0.00002 | 0.013 | 0.152 |
| | 69 | 11 (2) | 20: 10, 11, 18, **19**, **20** | 133 | 0.00086 | 0.091 | 0.101 |
| | 84 | 15 (6) | 20: 05, 08, 09, **10**, 11, 14, **15**, 16, 17, 19 | 104 | 0.00011 | 0.035 | 0.130 |

*Continued on next page*

| **Clan** | **Coda type** | **Recording year span** | **Recording years** | **Total codas** | **Line slope (s/year)** | **Mantel test matrix correlation coefficient** | **Mantel test p-value** |
|---|---|---|---|---|---|---|---|
| EC2 | 514* | 13 (13) | 20: **08**, 11, **12**, **16**, **18**, **19**, **20** | 540 | 0.00019 | 0.019 | 0.080 |
| | 515* | 13 (3) | 20: 08, 11, 12, 16, **18**, 19, **20** | 97 | 0.00011 | 0.019 | 0.306 |
| MED | 412* | 15 (10) | 20: **04**, 05, 06, 08, **13**, 14, 17, 18 | 148 | 0.00023 | 0.058 | ** |
| PO | 515* | 12 (5) | 19: 78, **85**, **87**, **89** | 168 | 0.00143 | 0.128 | * |
| REG | 64 | 11 (11) | 19: **85**, 87, 89, 91, **95** | 129 | 0.00342 | 0.193 | *** |
| | 73* | 11 (11) | 19: **85**, 87, 89, 91, **95** | 112 | 0.00436 | 0.278 | *** |
| | 713* | 11 (11) | 19: **85**, **87**, 89, 91, **95** | 159 | -0.00020 | -0.023 | 0.873 |
| | 83* | 11 (11) | 19: **85**, **87**, **89**, **91**, **95** | 176 | 0.00015 | 0.016 | 0.272 |
| | 810* | 11 (3) | 19: **85**, **87**, 89, 91, 95 | 101 | 0.00229 | 0.150 | ** |

### *5.4.3.2 – Modeling results*

Unless otherwise mentioned, the clan$_{\text{allyears}}$ and clan$_{\text{largeyears}}$ results were consistent (in terms of direction and significance), and only the clan$_{\text{allyears}}$ results are discussed. For the eight types that underwent additional analyses (i.e. the triangles in Figure S5.12), PC1 primarily represented total coda duration (explaining 50.6–98.7% of the variability) and PC2 typically represented specific ICI durations (explaining 0.5–30.4% of the variability; see panel (a) of Figures S5.13–S5.20). Duration and PC1 were thus largely redundant as dependent variables in our models, and only the duration results are discussed below unless the direction or significance of the PC1 results differed (see Table 5.2 for all model results).

*Table 5.2.* Clan-level modeling results for duration, PC1, and PC2. The ‘Coda type’ column gives numeric type codes; for type names, see Table S5.4 (EC1, MED) and Table S5.5 (PO, REG). See Table 5.1 for clan abbreviations. Stars in the ‘Coda type’ column denote identity coda types. The ‘Analysis’ column indicates if the results are from the clan$_{allyears}$ or clan$_{largeyears}$ analyses. For each dependent variable (duration, PC1, and PC2), the fixed effect (i.e. year) model coefficient and p-value are provided. P-value notation is as follows: 0=***, 0.001=**, 0.01=*. All values >0.01 are typed out, with significant p-values (≤0.05) bolded. If the repertoire random effect in the linear mixed-effects model was not significant and a linear model was run instead, the coefficient is italicized. Gray shading denotes regions included in the Pacific tree; no shading denotes regions included in the Atlantic/Mediterranean tree.

| **Clan** | **Coda type** | **Analysis** | **Total codas** | **Duration** | | **PC1** | | **PC2** | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | *Coefficient* | *p-value* | *Coefficient* | *p-value* | *Coefficient* | *p-value* |
| EC1 | 46 | All years | 191 | -0.002 | *** | -0.0014 | *** | *0.0001* | 0.311 |
| | | Large years | 145 | -0.001 | * | *-0.0009* | *** | *0.0001* | 0.739 |
| | 51* | All years | 1,972 | 0.001 | 0.142 | 0.0009 | 0.079 | 0.0005 | *** |
| | | Large years | 1,902 | 0.001 | 0.183 | 0.0009 | 0.115 | 0.0005 | ** |
| | 54* | All years | 215 | -0.001 | *** | 0.0006 | *** | 0.0001 | 0.658 |
| | | Large years | 147 | -0.001 | * | 0.0007 | **0.015** | 0.0004 | 0.141 |
| MED | 412* | All years | 148 | *-0.002* | **0.018** | *0.0009* | **0.020** | *0.0001* | 0.558 |
| | | Large years | 89 | -0.002 | 0.078 | 0.001 | 0.084 | 0.0001 | 0.763 |
| PO | 515* | All years | 168 | *-0.004* | **0.037** | *0.0021* | 0.052 | *-0.0010* | 0.058 |
| | | Large years | 165 | *-0.002* | 0.484 | *0.0007* | 0.588 | *-0.0020* | **0.015** |
| REG | 64 | All years | 129 | 0.013 | **0.020** | -0.0057 | **0.019** | *0.0001* | 0.540 |
| | | Large years | 78 | 0.013 | * | -0.0060 | * | *0.0001* | 0.566 |
| | 73* | All years | 112 | 0.020 | * | -0.0082 | * | *0.0001* | 0.332 |
| | | Large years | 58 | 0.022 | ** | -0.0090 | ** | *0.0002* | 0.278 |
| | 810* | All years | 101 | *0.004* | 0.357 | *0.0021* | 0.17265 | *0.0004* | **0.026** |
| | | Large years | 53 | *0.041* | 0.070 | *0.0090* | 0.120 | *0.0007* | 0.413 |

Coda type 46 is a non-identity coda made by EC1 clan whales that follows a ‘4D’ rhythmic pattern (Figure S5.4, Table S5.4). However, all but one coda of this type were made by whales belonging to social unit FU[26] (Tables 5.2), so any trends observed for type 46 codas may not be representative of clan-wide trends[27]. Modeling results showed that duration significantly decreased over time, while PC2 (which represented the ratio of the first and last ICI vs. the middle ICI) did not (Figure S5.21, Table 5.2). From 2005 to 2019, the duration of type 46 codas decreased by 0.002 s per year, resulting in a 0.03 s (–5.7%[28]) decrease over the 15-year period (Table 5.2).

The other non-identity coda with significant Mantel test results—type 64—is a ‘6R’ coda type made by Regular clan whales between 1985 and 1995 (Figure S5.6; Table S5.5). The modeling results were significant for duration but not PC2 (high PC2 codas have a relatively longer ICI1 vs. ICI5; Figure S5.18) (Table 5.2). The duration of coda type 64 increased by an average of 0.013 s per year, with a net increase of about 0.143 s (+10.1%) over an 11-year span (Table 5.2). However, violin plots of coda duration show that this increase was not linear or gradual over the examined timespan, but primarily occurred sometime between 1991 and 1995 (Figure S5.18).

Interestingly, a similar pattern was seen for Regular clan identity coda type 73 (a ‘7R’ coda). The model duration result was significant while the PC2 result was not (high PC2 codas have a relatively longer ICI2 vs. ICI1; Figure S5.19) (Table 5.2). Duration increased by 0.020 s per year between 1985 and 1995, with a net increase of 0.22 s (+13.2%) over the 11-year span (Table 5.2); once again, this increase was not gradual over time, but largely seems to have occurred between 1991 and 1995 (Figure S5.19).

Two EC1 clan identity codas had significant Mantel test results: type 51 (the most abundant coda type analyzed) and type 54 (Table 5.1). Both five-click coda types were recorded over 40-year spans (1981–2020) and have a ‘1+1+3’ rhythmic pattern (Table 5.1, Figure S5.3). For coda type 51, PC2 primarily represented the duration of the first

---

[26] Prior to 2008, social units F and U were considered separate, but by 2009 they met the definition of a single social unit (Konrad et al. 2018). They have been analyzed as a single social unit throughout this study.

[27] Indeed, coda type 46 is a non-identity coda at the clan level but is likely an identity coda at the unit-level. This was not explicitly tested here but is supported by Gero, Whitehead, et al. (2016).

[28] Percent change over time was calculated by dividing the net change (fixed effect coefficient * number of years) by the average duration or ICI ratio in the earliest year and multiplying by 100.

two ICIs (the long ICIs) vs. the last two ICIs (the short ICIs) (Figure S5.15). Only the model results for PC2 were significant (Table 5.2). This suggests that, at the clan-level, total coda duration did not change much over the timespans analyzed, but the ratio of long vs. short ICIs did. To look at this in more detail, we calculated the ratio of long ICIs (i.e. average of ICI1 and ICI2) to short ICIs (i.e. average of ICI3 and ICI4) for each coda of this type (Figure S5.22). We then ran additional models with ICI ratio as the dependent variable, year as a fixed effect, and repertoire as a random effect. The model results were significant, with the ratio increasing by 0.12 (+5.5%) over 40 years (Figure S5.22). With duration not significantly changing over time, this ratio increase likely resulted from the short ICIs decreasing in duration over time

In contrast, for EC1 identity coda type 54, the duration model result was significant but the PC2 result was not (Table 5.2). Here, PC2 represented the first and last ICI durations vs. the middle ICI durations (Figure S5.16). Between 1981 and 2020, the average duration of type 54 codas decreased by 0.001 s per year (Table 5.2), resulting in a 0.04 s (–4.8%) net decrease over the 40-year span (Figure S5.16). In other words, type 54 codas became faster over time.

Results were less conclusive for some of the other coda types when comparing the $\text{clan}_{\text{allyears}}$ and $\text{clan}_{\text{largeyears}}$ results. For coda type 412 (a ‘3+1’ patterned identity coda made by the Mediterranean clan; Figure S5.3), PC2 primarily represented the duration of the final ICI compared to the duration of the first two ICIs (Figure S5.14). Results from models showed that duration (but not PC2) significantly varied over time, with the average duration of type 412 codas decreasing between 2004 and 2018 by 0.002 s per year (Table 5.2). This translates to a net decrease of about 0.03 s (–4.4%) over the 15-year period (Figure S5.14). However, there was no significant difference in coda duration, PC1, or PC2 over time in the $\text{clan}_{\text{largeyears}}$ analysis (Figure S5.14, Table 5.2).

In the Regular clan, only the PC2 model result showed significant change for identity coda type 810 (an ‘8R’ coda type) in the $\text{clan}_{\text{allyears}}$ analysis (Table 5.2). PC2 roughly represented early vs. later ICIs (ICIs 1/2 vs. ICIs 6/7) in the coda (Figure S5.20). This result disappeared in the $\text{clan}_{\text{largeyears}}$ analysis, which left only codas spanning a three-year range (1985–1987; Table 5.2, Figure S5.20). Interestingly, visual inspection of violin plots of coda duration by year (Figure S5.20) suggest a similar trend of durational lengthening in 1995 as seen for coda types 64 and 73, but this was not confirmed statistically (i.e. the duration/PC1 model results weren't significant; Table 5.2).

Finally, one Plus-One clan coda type had a significant Mantel test result: identity coda type 515, which follows the '4+1' rhythmic pattern (Figure S5.5). For type 515, PC2 represents the duration of the final ICI compared to the three preceding ones (Figure S5.17). In the $\text{clan}_{\text{allyears}}$ analysis, the model with duration as the dependent variable was significant and suggested that the duration of type 515 codas shortened between 1978 and 1989 by 0.004 s per year (net decrease of 0.048 s, or –3.7%, over 12 years; Figure S5.17, Table 5.2). The models with PC1 and PC2 as the dependent variable approached but did not reach significance (p=0.052 and p=0.058, respectively; Table 5.2). However, this significant result was largely driven by three codas recorded in 1978. When that year was removed in the $\text{clan}_{\text{largeyears}}$ analysis, duration was no longer significant but PC2 was, with an increase in the final ICI in the final year, 1989 (Table 5.2; Figure S5.17).

## 5.5 – Discussion

The broad picture that emerges from our analyses is that sperm whale codas generally decrease in similarity over time, as evidenced by the Euclidean distance plot slopes increasing for 14 out of 18 examined types. However, these patterns are clearly nuanced and lend themselves to three additional conclusions: (1) temporal stability differs by coda type; (2) the magnitude of change over time can vary, and (3) coda durations can increase and decrease. Below, we discuss each result in more detail.

### *5.5.1 – Temporal stability differs by coda type*

Different coda types can clearly exhibit different degrees of stability over time, even within a clan. In the two clans with the largest number of examined types, EC1 and Regular, three out of nine and three out of five types had significantly changed over time, respectively. In contrast, neither of the EC2 clan's codas had significantly changed, and results from the Plus-One and Mediterranean clans were somewhat inconclusive. These patterns echo those seen in other taxa, including humans (Pagel et al., 2007), songbirds (Price & Lanyon, 2002), and frogs (Cocroft & Ryan, 1995): namely, that modifications can accumulate at different rates in different call types or call features. For example, in northern resident orcas, the N4 discrete call type changed over a 12–13-year period while the N9 discrete call type did not (Deecke et al., 2000). In southern resident orcas, the mean duration of 16 call types changed over 28 years while the mean duration of five call types remained stable (Wieland et al., 2010).

The Regular clan has two 7R (73 and 713) and two 8R (810 and 83) identity codas that were recorded between 1985 and 1995; types 73 and 810 had significant Mantel test results (i.e. showed evidence of change over time), but types 713 and 83 did not. Interestingly, types 713 and 83 had larger second largest year sample sizes than the two types that significantly changed, but even with small second largest year sample sizes, the effect sizes were quite large for types 73 and 810 (Figure S5.12). Similar variability was seen for EC1 clan codas that followed the 1+1+3 rhythmic pattern. Two 1+1+3 coda types (type 57, intermediate duration; type 511, shortest duration) did not show change over time, while two others (type 54, intermediate duration; type 51, longest duration) did; all four types were identity codas. Types 51, 57, and 511 had large second largest year sample sizes, suggesting that those analyses were appropriately powered (Figure S5.12). Type 54 had a small second largest year sample size, suggesting we should exercise more caution in interpreting the significant Mantel test result (Figure S5.12).

Collectively, these trends suggest that codas with the same rhythm but different tempos can exhibit different levels of stability over time and emphasize that tempo is likely an important feature of codas. As a reminder, the coda types themselves as determined by IDcall are not discrete (i.e. each coda is assigned a posterior probability of belonging to each type). While this could suggest that variation over time seen in some but not all codas of a given rhythmic pattern might simply be an artefact of some codas being poorly classified to types, this is unlikely to be the case here because we only analyzed codas with extremely high type classification certainty (i.e. had a type classification posterior probability ≥0.9).

### *5.5.2 – The magnitude of change over time can vary*

For coda types that have significantly changed over time, there is variability in the magnitude of change. Longer timespans were not always associated with larger amounts of annual or net change, and changes were not always linear over the timescales analyzed. Regular clan coda types increased in duration by 10.1% (type 64) and 13.2% (type 73) over 11 years, with much of that increase seemingly occurring between 1991 and 1995. In contrast, EC1 codas decreased in duration by just 4.8% over 40 years (type 54) and 5.7% over 15 years (type 46). For coda type 51, the ICI ratio decreased by 5.5% over 40 years. The annual changes in duration thus varied by clan and coda type but ranged from –0.38% (EC1 clan type 46) to +1.20% (Regular clan type 73) over the full timescales analyzed.

Similar rates of change have been seen in other cetacean vocalizations. In the spectral domain, the tonal frequencies of blue whale song decreased by 0.37–0.68% per year in seven populations (McDonald et al., 2009). In the temporal domain, southern resident orca discrete call durations changed anywhere from 0.39% to 7.8% per year (Wieland et al., 2010). It has been posited that rates of change can vary among different call types or call features if they convey different types of information (Williams et al., 2013), but we did not see significantly different rate trends for identity and non-identity codas (although our sample size was limited). For sperm whales, only so much within-type change is possible before a given type becomes a different type altogether. Future studies that quantify how much a coda type must change before becoming another type would help contextualize the rates of change observed here. Garland & McGregor (2020) clearly distinguish between rapid 'revolutions' and slower 'evolutions' in vocalizations. Humpback whale songs are an iconic example of the former, while the rates of change we observed for sperm whales suggest they generally fall in the latter category.

### *5.5.3 – Coda durations can increase and decrease*

In some taxa, like blue whales and fin whales, the call evolution documented to date has been strongly directional (McDonald et al., 2009; Weirathmueller et al., 2017). Sperm whale coda durations can decrease (e.g. EC1 coda type 54) or increase (e.g. Regular clan coda type 63), and there is at least preliminary evidence (albeit based on our very small sample size) that when coda durations do significantly change, clans are consistent in the direction of change. For example, two Regular clan coda types with different numbers of clicks (non-identity coda 64 and identity coda 73) showed similar trends in average coda duration over time, with stability from 1985 to 1991 but a significant increase in 1995. Identity coda type 810 had a small sample size (n=101) and inconsistent modeling results in the $\text{clan}_{\text{allyears}}$ and $\text{clan}_{\text{largeyears}}$ analyses, but violin plots of duration over the same 11-year timespan suggest an increase in duration in 1995 as well (although this was not confirmed statistically; Figure S5.20). In the EC1 clan, the durations of coda types 46 and 54 decreased over time (albeit over very different timespans of 15 vs. 40 years, respectively), but the duration of another coda type (51) did not significantly change.

Discrete call type durations were documented increasing and decreasing in southern resident orca calls, but increases were much more common (Wieland et al., 2010). After ruling out several alternative explanations (including maturational effects, seasonality, changes in group membership, and cultural drift), the researchers concluded that southern residents were generally lengthening their calls to combat increased vessel noise in their primary habitat (Wieland et al., 2010). In the present work on sperm whales, the observed trends generally seem to indicate nondirectional cultural drift (Deecke et al., 2000; Mundinger, 1980; Wieland et al., 2010). A potential exception is the Regular clan: the pattern of increased duration of 1995 codas seen across several click lengths and coda types could signify a cultural fad. Additional Galápagos Islands Regular clan coda recordings from 1995 or proceeding years could help investigate this further.

EC1 coda type 46 (which decreased in duration by a net of 0.03 s/5.7% over 15 years) is an interesting case. This 4D coda is not an identity coda for the EC1 clan as a whole, but may be a unit-specific coda for the best studied family of sperm whales in the world: unit FU (Gero, Whitehead, et al., 2016). While several members of FU have been recorded making this coda type, one female in particular—'Fingers'—seems to make the vast majority of type 46 codas (Gero, Whitehead, et al., 2016). In fact, of the 190 type 46 codas attributed to unit FU in the present study, Fingers was confirmed present or could have been present for 172 (90.6%) of them. We should thus exercise caution in saying that any trends in coda type 46 represent clan- or even unit-level trends, as they may largely represent the trends of a single individual.

### *5.5.4 – Limitations and future directions*

Despite starting with a prolific coda dataset (almost 40,000 codas), restricting our analyses to well-sampled types from well-sampled clans in small geographic regions, coupled with our extremely conservative inclusion criteria for codas and repertoires, greatly reduced the final sample size of analyzed codas (Table 5.1/S5.3). This made it difficult to distinguish whether certain apparent clan-level trends are actually social unit or even individual trends, as may be the case for EC1 coda type 46. This is more likely to be a potential issue in the Lesser Antilles dataset than in the Galápagos Islands dataset, because sperm whale encounters in the Lesser Antilles are often restricted to single, small, well-known social units, whereas Galápagos sperm whale encounters typically involve many more social units and whales (Whitehead et al., 2012). Future studies investigating coda type stability over time should target each level of sperm whale social structure.

While the codas analyzed here were recorded in various behavioral contexts, we have not explicitly controlled for behavioral context in our within-type analyses of temporal stability. Codas are produced in many social situations, ranging from two whales beginning a foraging dive to tens of whales interacting in high arousal social groups. Vocal output varies with behavioral context in other cetaceans with dialects, including short-finned pilot whales (Van Cise et al., 2018) and orcas (Filatova et al., 2013), as well as in other mammals (reviewed in Briefer 2012). In sperm whales, the effect of behavioral context on coda production specifically has received little research attention (but see Frantzis & Alexiadou, 2008). Future work quantifying whether certain coda features (e.g. type, duration) vary in different behavioral contexts will help address whether any of the temporal variation we documented here might reflect behavioral context at the time of recording.

It is worth reiterating that the timescales we examined are within the lifespan of an adult female sperm whale. The patterns (or absences) of change described here could be an artefact of the timescales examined. However, other cetacean species, including humpback whales (Garland et al., 2011), orcas (Deecke et al., 2000), and blue whales (McDonald et al., 2009), exhibit vocal evolution within individual lifespans of comparable durations to sperm whales. This suggests that, at least until longer term datasets exist, it is worthwhile to examine the available data.

### 5.6 – Conclusions

The goal of this research was to quantify if and how fine-scale structural parameters of sperm whale codas change over time within clans, and we showed that coda durations and ICI ratios can significantly change over time. Given that culture is a pervasive aspect of sperm whale life, we believe the most parsimonious explanation for the trends observed here is cultural drift, whereby random learning mistakes or innovations become fixed in a population. This fits with the within-clan patterns we observed over time, with some coda types staying stable, some changing in duration, and some changing in ICI ratio. A potential exception is the Regular clan, which may have undergone a cultural fad for lengthened codas in 1995, but additional codas are required to investigate this further.

Several key questions remain to be answered, including why some coda types change while others do not, and whether the variations we documented are perceptible and meaningful to the whales themselves. Whether a vocalization changes over time or not in a population can depend on its function, the selective forces acting on it, and the timescale examined. Given evidence that identity codas may be symbolic markers of clan identity (chapter 4), we hypothesized that identity coda types (n=12) would be more stable over time than non-identity coda types (n=6) across clans, but this was not confirmed statistically. Thus, we have no clear evidence that coda function, at least in terms of identity or non-identity designation, influences stability over time. This could be further investigated using more complex linear mixed-effects models, which include identity/non-identity coda designation, coda type, clan, year, and all possible interactions as fixed effects; repertoire as a random effect; and duration as the response variable.

Regarding the second question, most coda types changed by just a few milliseconds each year, which would be difficult for the human ear to detect. However, sperm whales possess a sophisticated auditory system with extremely high temporal resolution and may be able to perceive the changes documented here (Møhl et al., 2000). Additional research on sperm whale auditory discrimination abilities is necessary to assess this. Even if the whales can perceive these changes, they may not be meaningful (i.e. statistical significance may not translate to biological relevance); playback experiments can help address questions of perception and meaning in wild animals. The significant changes we documented, by virtue of our experimental design, were within the margin of variability for each coda type (i.e. the changes were not large enough for codas to be classified as a different type altogether). This suggests that coda types are more graded than has previously been suggested but does not detract from our ability to distinguish different clans based on the codas they produce. Thus, we believe that sperm whale clan dialects are stable enough over time (at the macro-scale) to be used in sperm whale conservation.

### 5.7 – Acknowledgements

We are grateful to Kristian Beedholm for creating and sharing CodaSorter and Koen de Reus for providing helpful feedback on a draft of this manuscript. We thank members of the Dominica Sperm Whale Project (http://www.thespermwhaleproject.org/), the Whitehead Lab, and the Balearic Sperm Whale Project (https://synergy.st-andrews.ac.uk/balearicspermwhales/), who collectively invested thousands of hours in fieldwork, audio analysis, and coda marking to build the long-term datasets used in this study.

# CHAPTER 6

# DISCUSSION

> “Knowledge is a rising tide. Lifts all boats, as it were.”
>
> —Victoria Aveyard

Dialect divergence among conspecifics in sympatry is rare in the animal kingdom[29] and, when it occurs, culture is often implicated (e.g. mountain white-crowned sparrows, *Zonotrichia leucophrys oriantha*, MacDougall-Shackleton & MacDougall-Shackleton 2001; orcas, Deecke et al., 2000, Filatova et al., 2015; humans, Hinskens et al., 2005; Hinskens, 2014). In sperm whales, there is a clear interplay between dialects and culture, and a major goal of my thesis was to better understand how vocalizations augment sociality over space and time in this highly cultural species. A dynamic picture emerges from this research, with different coda types and clans falling into different positions along a spectrum of stability and change.

## 6.1 – Summary and context of thesis findings

In chapter two, I showed that rhythm is pervasive in sperm whale (and other cetacean) vocalizations, regardless of behavioral context. This prevalence suggests that rhythm is a fundamental feature of sperm whale communication. Sperm whales surface and dive together but forage separately (Irvine et al., 2017), and the predictable, generally isochronous rhythm that characterizes echolocation likely helps whales keep track of one another and coordinate surface reunions. In the social domain, rhythm in codas occurs at multiple levels, with isochronous or heterochronous codas stitched together into isochronous coda bouts (Schulz et al., 2008). The ten clans that I examined in this thesis

---

[29] This is especially true compared to dialect divergence among heterospecifics in sympatry, where divergence usually evolves to prevent suboptimal mismating between closely-related species (e.g. green tree frogs, *Hyla* spp., Höbel and Gerhardt 2003; African tinkerbirds, *Pogoniulus* spp., Kirschel et al. 2009; southern wood crickets, *Gryllus* spp., Jang and Gerhardt 2006).

not only had clear preferences for specific coda types, but also often had rhythmic 'motifs' that characterized their dialects. For example, Plus-One clan codas typically have an extended pause before the final click, reminiscent of the Canadian tendency to add 'eh?' to the ends of sentences[30]. Previous work documented motifs in Regular, Plus-One, and Four-Plus clan dialects (Rendell & Whitehead, 2003b), but my thesis shows that rhythmic motifs are present in Slow Increasing, Rapid Increasing, Palindrome, EC2, and Mediterranean clan dialects as well.

In humans, different cultures have different rhythmic motifs in music and speech (Patel & Daniele, 2003; Widdess, 2013). For example, Jacoby and McDermott (2017) showed that perceptual priors (i.e. biases) on rhythm differed between American and Tsimané participants during a finger tapping task, such that certain priors were culture-specific. Priors can constrain cultural transmission because only signals with a high probability under the prior will be robustly produced and transmitted among individuals (Jacoby & McDermott, 2017). Agent-based models have shown that sperm whale clans likely originated from cultural transmission via biased social learning of codas, suggesting that there are priors on coda rhythm as well (Cantor et al., 2015). The different rhythmic motifs seen in the dialects of sympatric clans could indicate that, as in humans, those priors differ by sperm whale culture.

Most sperm whale coda studies to date have used the same categorical and continuous methods to divide codas into types and delineate sperm whale clans (Amorim et al., 2020; Cantor et al., 2016; Gero, Bøttcher, et al., 2016; Huijser et al., 2020; Rendell & Whitehead, 2003b). When I applied those methods to the Pacific Ocean coda dataset from chapter four, however, they did not perform well, likely because of the vast dataset size. Many codas were designated as noise and not assigned to a type during the categorical approach, and the dendrogram of coda repertoires that was built from the continuous approach output had extremely low bootstrap support. This prompted me to develop a new method for classifying codas into types and clustering repertoires into clans in chapter three.

---

[30] This fantastic analogy was made by my supervisor, Dr. Hal Whitehead, during a 2020 interview with Dal News.

The new method, IDcall, had two main sources of inspiration: fuzzy clustering and perceptual priors. The idea behind fuzzy clustering is that not all datasets can (or should) be divided into clean, discrete categories (Wadewitz et al., 2015). While codas do form clusters in PC space, some of those clusters are extremely small and dense, while others are more diffuse and graded (Figure 1.1). IDcall accounts for this by classifying calls into types using contaminated mixture models, which allow types to have different volumes, shapes, and orientations in multivariate space (Punzo & McNicholas, 2016). Additionally, while the categorical approach used in past studies treated coda types as bounded entities, the coda types produced by IDcall are not discrete: every call has some probability of belonging to every type. This affords the user greater control in future analytical steps, which can be tailored based on the specific research objectives. For example, I retained all codas when determining the number of Pacific clans in chapter four, because I wanted to capture the true variability of the dataset. In contrast, my chapter five analyses were restricted to codas with posterior probabilities of type classification ≥0.9, because I was looking for extremely fine-scale, within-type evidence of change over time.

If transmission is indeed limited to codas with high probability under culture-specific priors, perhaps focusing on the most frequently used coda types (i.e. those with the highest probability of being produced) could facilitate cleaner delineations of repertoires into clans? This is the approach taken in IDcall, where clans are only delineated if their constituent repertoires show uniformly high usage of a coda type that is rarely produced by any other clan. IDcall's strong performance on three distinct datasets shows that this underlying theory is applicable beyond sperm whales and can be used to detect putative biological structure in some insects and birds as well.

My analyses in chapter three confirm that humans (via IDcall) can detect the upper level of sperm whale sociality, the vocal clan, by looking at usage of codas with different rhythms and tempos. In this way, certain coda types have come to symbolize different sperm whale clans to us. Many researchers have hypothesized that certain coda types are used as symbolic markers of clan identity by the whales themselves as well (Cantor & Whitehead, 2013; Gero, Bøttcher, et al., 2016; Rendell & Whitehead, 2003b), but the data were only indicative prior to my thesis (Whitehead & Rendell, 2014).

Chapter four provides two pieces of evidence. The first is that, within clans, identity coda usage is more stable over geographic space than non-identity coda usage. This could signify that identity codas are the bedrock of the cultural dialect, while non-identity codas, which are more susceptible to spatial drift, potentially convey regional accents. The second is that, between clans, acoustic similarity based on identity (but not non-identity) coda usage decreases as clan spatial overlap increases. More spatially overlapped clans in the Pacific therefore appear to be more “marked” than less overlapped clans, which fits with patterns observed in symbolic markers of human ethnic groups (McElreath et al., 2003). Thus, while coda symbolic markers may not have facilitated the origination of clans (Cantor et al., 2015), my work suggests that they help maintain present-day clan divisions in areas of sympatry. These trends also suggest that identity and non-identity coda types have experienced different selective pressures over time.

I investigated coda temporal stability in more detail in chapter five and found that not all coda types are static entities. Most coda types tend to become less similar over time, and both identity and non-identity codas have significantly changed over decadal and multi-decadal timespans. This change usually manifests as increases or decreases in total coda duration over time. Within a clan, how can we rectify the stable usage (chapter four) but temporal lability (chapter five) of some identity coda types? If an identity coda changes within the lifespan of an individual sperm whale, how reliable is it as a marker of cultural identity?

Information can be encoded in multiple coda features, including frequency, rhythm, and tempo. Based on the results of my thesis, I believe that sperm whales primarily convey vocal clan (i.e. cultural) identity through the rhythmic patterning of codas. Theoretically, two sympatric clans could have identity codas with the same rhythmic pattern but different tempos; practically, I rarely observed this. When looking within individual clans, however, eight of the ten clans had at least two identity codas with the same rhythm but different tempos (e.g. EC1 1+1+3 codas, Mediterranean 3+1 codas, Plus-One 4+1 codas). Additionally, for most coda types that significantly changed over time, the rhythmic pattern of clicks remained the same, but the tempo changed. These results suggest that tempo may be less constrained than rhythm within sperm whales clans. Tempo variations likely contribute to the “fuzziness” of codas, with discrete rhythmic patterns that robustly convey cultural identity being the basis of more graded displays that could reflect arousal level, emotional state, and/or behavioral context (Cusano et al., 2021).

## 6.2 – Research implications

The results of my thesis suggest that sympatric clans are the norm rather than the exception and add to a large body of evidence that cultural differences run deep for sperm whales. These insights have significant implications for sperm whale psychology, cultural complexity, and conservation.

Symbolic markers have typically been heralded as a human idiosyncrasy (Boyd & Richerson, 1987; Cohen, 2012; McElreath et al., 2003), but my research provides empirical evidence that Pacific Ocean sperm whales use certain coda types as symbolic markers of clan identity (chapter four). Similar ongoing analyses using Atlantic Ocean and Indian Ocean sperm whale codas will shed light on whether this is a regional or global phenomenon. This has implications for sperm whale psychology, as it suggests the presence of a tag-based cooperation system in the species. In humans, there is evidence that accents were an evolutionarily viable tag that helped individuals maintain cooperative relationships in diffuse social networks (Cohen, 2012). Sperm whales exhibit cooperative behaviors within and between social units, including allocare, allonursing, and communal defense (Arnbom et al., 1987; Gero et al., 2009; Konrad et al., 2019; Whitehead, 1996). In areas of sympatry, these cooperative behaviors do not extend to whales from different clans. That between-clan acoustic similarity calculated from identity coda usage was modulated by clan spatial overlap (a proxy for degree of sympatry), but acoustic similarity calculated from non-identity coda usage was not, suggests that sperm whale identity codas may act as tags that help the whales maintain or reinforce cooperative relationships in regions with greater cultural diversity. This is in line with modelling work showing that symbolic markers can facilitate cooperative strategies that are resistant to exploitation and invasion by free-riders (Antal et al., 2009; Cohen, 2012; Ihara, 2011; Roberts, 2008).

Additionally, while human culture is often seen as the pinnacle of complexity, many animal cultures are complex as well (Schuppli & van Schaik, 2019). My thesis adds to our understanding of sperm whale cultural complexity, showing that different clans operate over different ecological scales; use different identity codas; and exhibit different within-coda type structural stability over time. The variation in clan scale is particularly striking, with some Pacific Ocean clans present in very small regions (e.g. the Plus-One clan) while others span the ocean basin (e.g. the Short clan). This pattern—different sperm whale cultural groups exhibiting sometimes extensive differences in distribution—mirrors what is seen in human ethnic groups (Ahlerup & Olsson, 2012; Nettle, 2009) and orca ecotypes (de Bruyn et al., 2013). Coupled with recent work by Vachon et al. (submitted) in the eastern Caribbean, the variability in vocal clan distribution in the Pacific suggests that the common characterization of sperm whales as 'ocean nomads' (Mizroch & Rice, 2013) may not always be accurate.

In a similar vein, my work in chapter five emphasizes that sperm whale social communication is dynamic. Clans not only exhibit differences in coda type preferences but also in within-type structural stability over time, with some coda durations increasing, decreasing, or not changing over various time scales. This work contrasts with the traditional perception of codas as static, discrete entities. It also adds sperm whales to a growing list of cetaceans whose vocalizations can change within individual lifespans (e.g. humpback whales, Garland et al. 2011; orcas, Deecke et al. 2000; blue whales, McDonald et al. 2009).

Aside from the Mediterranean subpopulation, which is classified as 'endangered' by the International Union for Conservation of Nature (IUCN), the global population of sperm whales is considered 'vulnerable' with unknown population trends (Taylor et al., 2019). Given that cultural differences can affect sperm whale fitness (Marcoux, Rendell, et al., 2007; Whitehead & Rendell, 2004), a growing contingent of researchers is advocating for incorporating culture into sperm whale conservation (e.g. Vachon et al., submitted; Whitehead et al., 2004; Bermant et al., 2019). This reflects a broader movement championed by those who study non-human animal culture to factor culture into conservation (Allen, 2019; Brakes et al., 2019, 2021; Crates et al., 2021; Ryan, 2006; Whitehead et al., 2004; Whiten, 2021).

Populations of animals are often divided into evolutionarily significant units (ESUs) for conservation (Ryder, 1986). These ESUs have historically been designated when subpopulations are genetically, geographically, or phenotypically distinct (Moritz, 1994). As our understanding of the reach and significance of culture has grown, the idea that culture should be incorporated into ESUs has gained traction as well, since culture can clearly divide population and is an evolutionarily significant process (Allen, 2019; Brakes et al., 2019, 2021; Ryan, 2006; United Nations Environmental Program, 2014; Whitehead et al., 2004). What would a culturally sensitive conservation framework look like for sperm whales, and how does my research contribute to it?

My thesis strongly suggests that vocal clans are an appropriate basis for designating ESUs for sperm whales, echoing previous work (Vachon et al., submitted; Whitehead et al., 2004; Convention on Migratory Species, 2017; Bermant et al., 2019; Brakes et al., 2021). It also provides an improved tool—IDcall—that can reliably ascertain vocal clan membership from sperm whale recordings (chapter three), which is essential to subdivide sperm whales at the vocal clan level. Using IDcall, we have learned that at least seven sperm whale clans inhabit the Pacific Ocean (chapter four) and another seven inhabit the North Atlantic Ocean (Shane Gero, personal communication). Additional analyses are ongoing for the Indian Ocean (Chris Johnson, Curtin University) and Mediterranean Sea (myself). This highly collaborative work will give us an unprecedented understanding of the global extent and distribution of sperm whale cultural diversity. With that understanding, scientists around the world can conduct more targeted research on the clans that inhabit local waters, and better elucidate whether different clans face different challenges (e.g. the differential foraging success of Plus-One and Regular clan whales off the Galápagos Islands during El Niño and La Niña years; Whitehead & Rendell, 2004).

To illustrate this, consider Dominican sperm whales, which have been regularly studied since the Dominica Sperm Whale Project was established in 2005 by Shane Gero and Hal Whitehead. For the first decade of the project, it was thought that all eastern Caribbean sperm whales belonged to a single vocal clan (Gero, 2012), in contrast to the sympatric clans found in the eastern tropical Pacific and off Japan (Amano et al., 2014; Whitehead et al., 2012). This was disproved in 2016, when analyses of recordings from rarely encountered social units provided concrete evidence of two clans, EC1 and EC2 (Gero, Bøttcher, et al., 2016). The EC2 clan clearly spent less time in Dominican waters, making up just 3% of photo-identified encounters between 2005 and 2019 (Vachon et al., submitted). This was a major contributor towards the impetus for expanded sperm whale surveys of the Lesser Antilles island chain in 2019 and 2020 which, among other things, documented a putative third clan (EC3) and showed that EC1 and EC2 whales have different distributions (Vachon et al., submitted). Equipped with this new knowledge of sperm whale clan structure and distribution in the eastern Caribbean, we can now feasibly measure clan population trajectories separately. We are also better equipped to detect whether different clans face different pressures that lead to fitness consequences (e.g. if clan entanglement rates differ and are correlated with fishing pressure off certain islands).

If international (e.g. IUCN, International Whaling Commission) and national (e.g. Committee on the Status of Endangered Wildlife in Canada, National Oceanic and Atmospheric Administration) government agencies can be convinced that sperm whale clans are the right ‘unit to conserve’ for the species, new legislative and financial opportunities that can improve conservation efforts may become available. For example, different clans could have different IUCN Red List Categories (if warranted), including endangered (which may trigger additional legal protections) and data deficient (which can funnel money towards additional research). All the available evidence suggests that cultural divisions matter to sperm whales, which means those divisions should be accounted for in conservation strategies. In this species and others, maintaining cultural diversity is likely just as important as maintaining genetic diversity when it comes to ensuring a healthy population of animals (Brakes et al., 2021).

### 6.3 – Looking back: a brief ode to archival data

My thesis relied almost entirely on acoustic data I did not collect. This enabled me to ask questions that would have been impossible without the cross-species (chapters two and three), spatial (chapter four), and temporal (chapter five) scales of data I had access to. It also forced me to confront the many challenges that come with working with archival data. That meant learning how to digitize reel-to-reel tapes recorded in the 1980s, figuring out how to turn corrupted Kay sonograph files into functional .wav files, and painstakingly scouring old logbooks and datasheets for any mention of codas. But it also meant that I, who study sperm whale codas over time and space, got to travel through the times and spaces of the researchers before me. I saw logbook pages peppered with drawings from bored crew members on whale-less days; found dried bits of sperm whale skin sandwiched between old datasheets; heard earnestly professional graduate student voiceovers at the starts of recordings devolve into peals of laughter. There were moments of frustration and moments of elation, but most often just quiet moments of dedicated individuals studying an enigmatic species. I am grateful to the giants on whose shoulders I stood[31], and for the privilege of working with their hard-earned data. There is so much left to find in archival datasets if you are willing to look.

### 6.4 – Looking ahead: new frontiers

Unsurprisingly, my thesis unearthed more questions than answers, and I am excited to see what the next years hold for sperm whale research. Results from my Pacific analyses suggest that sympatric clans are the norm rather than the exception for sperm whales, with the number of detected clans in a region generally increasing as the number of recorded codas and repertoires increases. Our list of known clans will likely continue to grow as work in understudied and remote regions continues; bottom-mounted hydrophone arrays and autonomous underwater gliders can expand the spatial and temporal scales of such research and make it more feasible.

Simultaneously, we urgently need to develop automated or semi-automated methods that can process the barrage of acoustic data that comes from modern recording technology. This includes methods to detect sperm whale clicks, discriminate codas, and—the holy grail—mark codas from raw recordings. The time investment currently required for this process is immense and limiting. For example, despite having access to acoustic recordings, I was unable to include several Pacific locations (e.g. Saipan, Wake Atoll, Gwaii Haanas) in my chapter four analyses because there was no way to efficiently process the data. Recent collaborations between sperm whale biologists and machine learning experts suggest that automated methods are on the horizon (Andreas et al., 2021;

---

[31] "If I have seen further it is by standing on the shoulders of Giants." (Isaac Newton, 1965)

Bermant et al., 2019), and I encourage the researchers developing them to embrace the "fuzziness" of codas in their work.

Playback experiments are, in my mind, the most concrete way to validate that sperm whales use certain coda types as symbolic markers of clan identity. Such experiments are notoriously difficult to conduct on free-ranging, deep-diving cetaceans, but the recently established Cetacean Translation Initiative is working to make sperm whale playbacks more feasible (Andreas et al., 2021). In the meantime, additional research that examines the structure of identity and non-identity codas in more detail and contextualizes the results in clan dynamics, such as degree of clan sympatry, will help us better understand the symbolic potential of identity codas.[32]

Research on animal vocal rhythms has burgeoned over the past decade and has helped us better understand both how animals use rhythm in their daily lives and the evolutionary trajectory of human rhythmic abilities (Benichov et al., 2016; de Reus et al., 2021; Hartbauer & Roemer, 2016; Norton & Scharff, 2016; Ravignani et al., 2014, 2016; ten Cate et al., 2016). My goal in chapter two was to usher cetaceans into the fray by showing that their vocal communication systems are rife with rhythm. Despite the many examples I found, few studies explicitly sought to better understand vocal rhythm. Future sperm whale work that quantifies rhythm in coda bouts (e.g. de Reus et al., 2021), investigates how behavioral and/or social context impacts rhythm production (e.g. Zahavi & Zahavi, 1999), and looks for evidence of categorical rhythms (e.g. Roeske et al., 2020; De Gregorio et al., 2021) will enhance our understanding of how sperm whales make sense of their world.[33]

### 6.5 – Conclusions

My doctoral research has helped shade in our portrait of sperm whales, further illustrating the dynamism of this species, as well as the form and function of their cultural dialects. Through this work, I have shown that rhythm is a fundamental feature of sperm whale vocal communication; that sperm whales, like humans, may use certain coda types as symbolic markers of cultural identity; and that different coda types exhibit different

---

[32] This is work that my lab mate, Félicia Vachon, is undertaking as part of her Ph.D.

[33] My lab mate, Ana Eguiguren, is conducting research on rhythm in coda bouts and the effect of context on rhythm production as part of her Ph.D.

levels of stability over time. This work emphasizes the need for a global, multi-cultural approach to studying and conserving a global, multi-cultural species. Sperm whales clearly do not abide by the arbitrary boundaries we draw around them; it is time we start respecting the boundaries they draw around themselves.

# REFERENCES

Aguilar de Soto, N., Visser, F., Tyack, P. L., Alcazar, J., Ruxton, G., Arranz, P., Madsen, P. T., & Johnson, M. (2020). Fear of killer whales drives extreme synchrony in deep diving beaked whales. *Scientific Reports*, *10*(1), 1–9. https://doi.org/10.1038/s41598-019-55911-3

Aguilar Soto, N., Johnson, M. P., Madsen, P. T., Diaz, F., Dominguez, I., Brito, A., & Tyack, P. (2008). Cheetahs of the deep sea: Deep foraging sprints in short-finned pilot whales off Tenerife (Canary Islands). *Journal of Animal Ecology*, *77*(5), 936–947. https://doi.org/10.1111/j.1365-2656.2008.01393.x

Ahlerup, P., & Olsson, O. (2012). The roots of ethnic diversity. *Journal of Economic Growth*, *17*(2), 71–102. https://doi.org/10.1007/s10887-011-9075-0

Ainslie, M. A., & McColm, J. G. (1998). A simplified formula for viscous and chemical absorption in sea water. *The Journal of the Acoustical Society of America*, *103*(3), 1671–1672. https://doi.org/10.1121/1.421258

Allen, J. A. (2019). Community through culture: From insects to whales. *BioEssays*, *41*(11), 1–8. https://doi.org/10.1002/bies.201900060

Amano, M., Kourogi, A., Aoki, K., Yoshioka, M., & Mori, K. (2014). Differences in sperm whale codas between two waters off Japan: possible geographic separation of vocal clans. *Journal of Mammalogy*, *95*(1), 169–175. https://doi.org/10.1644/13-mamm-a-172

Amante, C., & Eakins, B. W. (2009). *ETOPO1 arc-minute global relief model: procedures, data sources and analysis*.

Amorim, T. O. S., Rendell, L., Di Tullio, J., Secchi, E. R., Castro, F. R., & Andriolo, A. (2020). Coda repertoire and vocal clans of sperm whales in the western Atlantic Ocean. *Deep-Sea Research Part I: Oceanographic Research Papers*, *160*, 103254. https://doi.org/10.1016/j.dsr.2020.103254

André, M., & Kamminga, C. (2000). Rhythmic dimension in the echolocation click trains of sperm whales: A possible function of identification and communication. *Journal of the Marine Biological Association of the United Kingdom*, *80*(1), 163–169. https://doi.org/10.1017/S002531549900168X

Andreas, J., Beguš, G., Bronstein, M. M., Diamant, R., Delaney, D., Gero, S., Goldwasser, S., Gruber, D. F., de Haas, S., Malkin, P., Payne, R., Petri, G., Rus, D., Sharma, P., Tchernov, D., Tønnesen, P., Torralba, A., Vogt, D., & Wood, R. J. (2021). Cetacean Translation Initiative: a roadmap to deciphering the communication of sperm whales. *ArXiv*, 1–29. http://arxiv.org/abs/2104.08614

Anichini, M., De Heer Kloots, M., & Ravignani, A. (2020). Interactive rhythms in the wild, in the brain, and in silico. *Canadian Journal of Experimental Psychology*, *74*(3), 170–175. https://doi.org/10.1037/cep0000224

Ankerst, M., Breunig, M. M., Kriegel, H. P., & Sander, J. (1999). OPTICS: Ordering points to identify the clustering structure. *ACM SIGMOD Record*, *28*(2), 49–60. https://doi.org/10.1145/304181.304187

Anshel, A., & Kipper, D. A. (1988). The influence of group singing on trust and cooperation. *Journal of Music Therapy*, *25*(3), 145–155. https://doi.org/10.1093/jmt/25.3.145

Antal, T., Ohtsuki, H., Wakeley, J., Taylor, P. D., & Nowak, M. A. (2009). Evolution of cooperation by phenotypic similarity. *Proceedings of the National Academy of Sciences*, *106*(21), 8597–8600. https://doi.org/10.1073/pnas.0902528106

Antunes, R., Schulz, T., Gero, S., Whitehead, H., Gordon, J., & Rendell, L. (2011). Individually distinctive acoustic features in sperm whale codas. *Animal Behaviour*, *81*(4), 723–730. https://doi.org/10.1016/j.anbehav.2010.12.019

Arnbom, T., Papastavrou, V., Weilgart, L. S., & Whitehead, H. (1987). Sperm whales react to an attack by killer whales. *Journal of Mammalogy*, *68*(2), 450–453. https://doi.org/10.2307/1381497

Arriaga, G., & Jarvis, E. D. (2013). Mouse vocal communication system: are ultrasounds learned or innate? *Brain and Language*, *124*(1), 96–116. https://doi.org/10.1016/j.bandl.2012.10.002

Arriaga, G., Zhou, E. P., & Jarvis, E. D. (2012). Of mice, birds, and men: the mouse ultrasonic song system has some features similar to humans and song-learning birds. *Plos One*, *7*(10), e46610. https://doi.org/10.1371/journal.pone.0046610

Atema, J. (1995). Chemical signals in the marine environment: dispersal, detection, and temporal signal analysis. *Proceedings of the National Academy of Sciences*, *92*(1), 62–66. https://doi.org/10.1073/pnas.92.1.62

Backus, R. H., & Schevill, W. E. (1966). Physeter clicks. In *Whales, dolphins, and porpoises* (pp. 510–527). University of California Press Berkeley. https://doi.org/https://doi.org/10.1525/9780520321373-030

Baker, C. S. (1985). *The population structure and social organization of humpback whales (Megaptera novaeangliae) in the central and eastern North Pacific*. University of Hawaii, Honolulu.

Balaban, E. (1988). Cultural and genetic variation in swamp sparrows (Melospiza georgiana): II. Behavioral salience of geographic song variants. *Behaviour*, 292–322. https://doi.org/10.1163/156853988x00043

Bannister, J. L., Josephson, E. A., Reeves, R. R., & Smith, T. D. (2012). There she blew! Yankee sperm whaling grounds, 1760–1920. In *Oceans Past* (pp. 135–158). Routledge.

Barker, A. J., Veviurko, G., Bennett, N. C., Hart, D. W., Mograby, L., & Lewin, G. R. (2021). Cultural transmission of vocal dialect in the naked mole-rat. *Science*, *371*(6528), 503–507.

Barkley, Y., Oleson, E. M., Oswald, J. N., & Franklin, E. C. (2019). Whistle classification of sympatric false killer whale populations in Hawaiian waters yields low accuracy rates. *Frontiers in Marine Science*, *6*, 645.

Baudry, J.-P., Raftery, A. E., Celeux, G., Lo, K., & Gottardo, R. (2010). Combining mixture components for clustering. *Journal of Computational and Graphical Statistics*, *19*(2), 332–353. https://doi.org/10.1198/jcgs.2010.08111

Beale, T. (1839). *The natural history of the sperm whale*. J. Van Voorst.

Belanger, R. M., & Corkum, L. D. (2009). Review of aquatic sex pheromones and chemical communication in anurans. *Journal of Herpetology*, *43*(2), 184–191.

Bell, A. V, Richerson, P. J., & McElreath, R. (2009). Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences*, *106*(42), 17671–17674.

Benedict, L., & Bowie, R. C. K. (2009). Macrogeographical variation in the song of a widely distributed African warbler. *Biology Letters*, *5*(4), 484–487.

Benichov, J. I., Globerson, E., & Tchernichovski, O. (2016). Finding the beat: from socially coordinated vocalizations in songbirds to rhythmic entrainment in humans. *Frontiers in Human Neuroscience*, *10*, 255.

Bermant, P. C., Bronstein, M. M., Wood, R. J., Gero, S., & Gruber, D. F. (2019). Deep machine learning techniques for the detection and classification of sperm whale bioacoustics. *Scientific Reports*, *9*(1), 1–10.

Bermeitinger, C., & Frings, C. (2015). Rhythm and attention: Does the beat position of a visual or auditory regular pulse modulate T2 detection in the attentional blink? *Frontiers in Psychology*, *6*, 1847. https://doi.org/10.3389/fpsyg.2015.01847

Best, P. B. (1979). Social organization in sperm whales, Physeter macrocephalus. In *Behavior of marine animals* (pp. 227–289). Springer.

Biernacki, C., Celeux, G., & Govaert, G. (2000). Assessing a mixture model for clustering with the Integrated Classification Likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(7), 719–725. https://doi.org/https://doi.org/10.1109/34.865189

Bispham, J. (2006). Rhythm in music: What is it? Who has it? And why? *Music Perception*, *24*(2), 125–134. https://doi.org/10.1525/mp.2006.24.2.125

Bivand, R., Rundel, C., Pebesma, E., Stuetz, R., Hufthammer, K. O., & Bivand, M. R. (2017). Package 'rgeos.' *The Comprehensive R Archive Network (CRAN)*.

Blumstein, D. T., & Armitage, K. B. (1997). Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. *The American Naturalist*, *150*(2), 179–200.

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: a comparison of three non-human primate species. *Frontiers in Psychology*, *4*, 390. https://doi.org/10.3389/fpsyg.2013.00390

Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S. J., Alekseyenko, A. V., Drummon, A. J., Gray, R. D., Suchard, M. A., & Atkinson, Q. D. (2012). Mapping the origins and expansion of the Indo-European language family. *Science*, *337*(6097), 957–960. https://doi.org/10.4159/harvard.9780674333987.c22

Bouwer, F. L., Nityananda, V., Rouse, A. A., & ten Cate, C. (2021). Rhythmic abilities in humans and non-human animals: a review and recommendations from a methodological perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1835), 20200335. https://doi.org/10.1098/rstb.2020.0335

Bowling, D. L., Herbst, C. T., & Fitch, W. T. (2013). Social origins of rhythm? Synchrony and temporal regularity in human vocalization. *PLoS ONE*, *8*(11), e80402. https://doi.org/10.1371/journal.pone.0080402

Boyd, R., & Richerson, P. J. (1987). The the evolution of ethnic markers. *Cultural Anthropology*, *21*(1), 65–79. https://doi.org/10.1525/can.1987.2.1.02a00070

Brakes, P., Carroll, E. L., Dall, S. R. X., Keith, S. A., McGregor, P. K., Mesnick, S. L., Noad, M. J., Rendell, L., Robbins, M. M., Rutz, C., Thornton, A., Whiten, A., Whiting, M. J., Aplin, L. M., Bearhop, S., Ciucci, P., Fishlock, V., Ford, J. K. B., Notarbartolo di Sciara, G., … Garland, E. C. (2021). A deepening understanding of animal culture suggests lessons for conservation. *Proceedings of the Royal Society B*, *288*(1949), 20202718. https://doi.org/10.1098/rspb.2020.2718

Brakes, P., Dall, S. R. X., Aplin, L. M., Bearhop, S., Carroll, E. L., Ciucci, P., Fishlock, V., Ford, J. K. B., Garland, E. C., Keith, S. A., McGregor, P. K., Mesnick, S. L., Noad, M. J., Notarbartolo di Sciara, G., Robbins, M. M., Simmonds, M. P., Spina, F., Thornton, A., Wade, P. R., … Rutz, C. (2019). Animal cultures matter for conservation. *Science*, *363*(6431), 1032–1034. https://doi.org/10.1126/science.aaw3557

Bresnahan, A. (2019). Dance rhythm. In P. Cheyne, A. Hamilton, & M. Paddison (Eds.), *The Philosophy of Rhythm: Aesthetics, Music, Poetics* (pp. 91–98). Oxford University Press. https://doi.org/https://doi.org/10.1093/oso/9780199347773.003.0006

Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. *Journal of Zoology*, *288*(1), 1–20.

Brumm, H., & Slater, P. J. B. (2006). Ambient noise, motor fatigue, and serial redundancy in chaffinch song. *Behavioral Ecology and Sociobiology*, *60*(4), 475–481. https://doi.org/10.1007/s00265-006-0188-y

Burchardt, L. S., & Knörnschild, M. (2020). Comparison of methods for rhythm analysis of complex animals' acoustic signals. *PLoS Computational Biology*, *16*(4). https://doi.org/10.1371/journal.pcbi.1007755

Burchardt, L. S., Norton, P., Behr, O., Scharff, C., & Knörnschild, M. (2019). General isochronous rhythm in echolocation calls and social vocalizations of the bat Saccopteryx bilineata. *Royal Society Open Science*, *6*(1). https://doi.org/10.1098/rsos.181076

Burchardt, L. S., Picciulin, M., Parmentier, E., & Bolgan, M. (2021). A primer on rhythm quantification for fish sounds: a Mediterranean case study. *Royal Society Open Science*, *8*(9), 210494.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, *3*(1), 1–27. https://doi.org/https://doi.org/10.1080/03610927408827101

Cantor, Maurício, Shoemaker, L. G., Cabral, R. B., Flores, C. O., Varga, M., & Whitehead, H. (2015). Multilevel animal societies can emerge from cultural transmission. *Nature Communications*, *6*(1), 1–10. https://doi.org/10.1038/ncomms9091

Cantor, Mauricio, & Whitehead, H. (2013). The interplay between social networks and culture: Theoretically and among Whales and Dolphins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1618). https://doi.org/10.1098/rstb.2012.0340

Cantor, Maurício, & Whitehead, H. (2015). How does social behavior differ among sperm whale clans? *Marine Mammal Science*, *31*(4), 1275–1290. https://doi.org/10.1111/mms.12218

Cantor, Mauricio, Whitehead, H., Gero, S., & Rendell, L. (2016). Cultural turnover among Galápagos sperm whales. *Royal Society Open Science*, *3*(10), 160615. https://doi.org/10.1098/rsos.160615

Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, *50*(11), 2652–2658. https://doi.org/10.1016/j.neuropsychologia.2012.07.018

Cerchio, S., Andrianantenaina, B., Lindsay, A., Rekdahl, M., Andrianarivelo, N., & Rasoloarijao, T. (2015). Omura's whales (Balaenoptera omurai) off northwest Madagascar: Ecology, behaviour and conservation needs. *Royal Society Open Science*, *2*(10), 150301. https://doi.org/10.1098/rsos.150301

Cerchio, S., & Dahlheim, M. (2001). Variation in feeding vocalizations of humpback whales Megaptera novaeangliae from southeast Alaska. *Bioacoustics*, *11*(4), 277–295. https://doi.org/10.1080/09524622.2001.9753468

Cerchio, S., Dorning, S., Andrianantenaina, B., & Cholewiak, D. (2017). A first description of rhythmic song in Omura's whale (Balaenoptera omurai). *The Journal of the Acoustical Society of America*, *141*(5), 3544.

Chase, O. (2015). *Shipwreck of the whaleship Essex*. Random House.

Christal, J., Whitehead, H., & Lettevall, E. (1998). Sperm whale social units: variation and change. *Canadian Journal of Zoology*, *76*(8), 1431–1440. https://doi.org/10.1139/cjz-76-8-1431

Clark, C. W. (1998). Whale voices from the deep: Temporal patterns and signal structures as adaptations for living in an acoustic medium. *The Journal of the Acoustical Society of America*, *103*(5), 2957–2957. https://doi.org/10.1121/1.422335

Clarke, M. R. (1978). Structure and proportions of the spermaceti organ in the sperm whale. *Journal of the Marine Biological Association of the United Kingdom*, *58*(1), 1–17.

Clarke, M. R. (1980). Cephalopoda in the diet of sperm whales of the southern hemisphere and their bearing on sperm whale biology. *Discovery Reports*, *37*, 1–324.

Cocroft, R. B., & Ryan, M. J. (1995). Patterns of advertisement call evolution in toads and chorus frogs. *Animal Behaviour*, *49*(2), 283–303.

Cohen, E. (2012). The evolution of tag-based cooperation in humans: The case for accent. *Current Anthropology*, *53*(5), 588–616. https://doi.org/10.1086/667654

Convention on Migratory Species. (2017). Concerted action for sperm whales (Physeter macrocephalus) of the eastern tropical Pacific. *UNEP/CMS/Concerted Action 12.2*. www.cms.int/en/document/ concerted-action-sperm-whales-physeter-macrocephalus-eastern-tropical-pacific

Coss, R. G., McCowan, B., & Ramakrishnan, U. (2007). Threat-related acoustical differences in alarm calls by wild bonnet macaques (Macaca radiata) elicited by python and leopard models. *Ethology*, *113*(4), 352–367.

Crance, J. L., Berchok, C. L., Wright, D. L., Brewer, A. M., & Woodrich, D. F. (2019). Song production by the North Pacific right whale, Eubalaena japonica. *The Journal of the Acoustical Society of America*, *145*(6), 3467–3479. https://doi.org/10.1121/1.5111338

Cranford, T. W. (1999). The sperm whale's nose: Sexual selection on a grand scale? *Marine Mammal Science*, *15*(4), 1133–1157.

Crates, R., Langmore, N., Ranjard, L., Stojanovic, D., Rayner, L., Ingwersen, D., & Heinsohn, R. (2021). Loss of vocal culture and fitness costs in a critically endangered songbird. *Proceedings of the Royal Society B*, *288*(1947), 20210225.

Cressey, D. (2015). World's whaling slaughter tallied. *Nature News*, *519*(7542), 140.

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees produce group-specific calls: A case for vocal learning? *Ethology*, *110*(3), 221–243. https://doi.org/10.1111/j.1439-0310.2004.00968.x

Croll, D. A., Clark, C. W., Acevedo, A., Tershy, B., Flores, S., Gedamke, J., & Urban, J. (2002). Only male fin whales sing loud songs. *Nature*, *417*(6891), 809. https://doi.org/10.1038/417809a

Currie, T. E., & Mace, R. (2009). Political complexity predicts the spread of ethnolinguistic groups. *Proceedings of the National Academy of Sciences*, *106*(18), 7339–7344.

Currie, T. E., & Mace, R. (2012). The evolution of ethnolinguistic diversity. *Advances in Complex Systems*, *15*(1–2), 1–20. https://doi.org/10.1142/S0219525911003372

Currie, T. E., Meade, A., Guillon, M., & Mace, R. (2013). Cultural phylogeography of the Bantu Languages of sub-Saharan Africa. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1762). https://doi.org/10.1098/rspb.2013.0695

Cusano, D. A., Conger, L. A., Van Parijs, S. M., & Parks, S. E. (2019). Implementing conservation measures for the North Atlantic right whale: considering the behavioral ontogeny of mother-calf pairs. *Animal Conservation*, *22*(3), 228–237. https://doi.org/10.1111/acv.12457

Cusano, Dana A., Noad, M. J., & Dunlop, R. A. (2021). Fuzzy clustering as a tool to differentiate between discrete and graded call types. *JASA Express Letters*, *1*(6), 61201.

Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua*, *92*, 81–104. https://doi.org/10.1016/0024-3841(94)90338-7

De Boer, B. (2010). Modelling vocal anatomy's significant effect on speech. *Journal of Evolutionary Psychology*, *8*(4), 351–366.

de Bruyn, P. J. N., Tosh, C. A., & Terauds, A. (2013). Killer whale ecotypes: Is there a global model? *Biological Reviews*, *88*(1), 62–80. https://doi.org/10.1111/j.1469-185X.2012.00239.x

De Gregorio, C., Valente, D., Raimondi, T., Torti, V., Miaretsoa, L., Friard, O., Giacoma, C., Ravignani, A., & Gamba, M. (2021). Categorical rhythms in a singing primate. *Current Biology*.

de Reus, K., Masayo, S., Marianna, A., Gamba, M., de Heer Kloots, M., Miriam, L., Bruno, J. H., Laurel, T., & Ravignani, A. (2021). Rhythm in dyadic interactions. *Philosophical Transactions of the Royal Society of London, Series B*, *376*, 20200337. https://doi.org/10.1098/rstb.2020.0337

De Waal, F. B. M. (1988). The communicative repertoire of captive bonobos (Pan paniscus), compared to that of chimpanzees. *Behaviour*, *106*(3–4), 183–251. https://doi.org/10.1163/156853988x00269

Deecke, V.B., Ford, J. K. B., & Spong, P. (2000). Dialect change in resident killer whales: implications for vocal learning and cultural transmission. *Animal Behaviour*, *60*(August 2016), 629–638. https://doi.org/10.1006/anbe.2000.1454

Deecke, Volker B. (2006). Studying marine mammal cognition in the wild: a review of four decades of playback experiments. *Aquatic Mammals*, *32*(4), 461–482.

Delarue, J., Laurinolli, M., & Martin, B. (2009). Bowhead whale (Balaena mysticetus) songs in the Chukchi Sea between October 2007 and May 2008. *The Journal of the Acoustical Society of America*, *126*(6), 3319–3328. https://doi.org/10.1121/1.3257201

Delarue, J., Todd, S. K., Van Parijs, S. M., & Di Iorio, L. (2009). Geographic variation in Northwest Atlantic fin whale (Balaenoptera physalus) song: Implications for stock structure assessment. *The Journal of the Acoustical Society of America*, *125*(3), 1774–1782. https://doi.org/10.1121/1.3068454

Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature*, *266*(5604), 718–719. https://doi.org/10.1038/266718a0

Di Paolo, E. A. (2000). Behavioral coordination, structural congruence and entrainment in a simulation of acoustically coupled agents. *Adaptive Behavior*, *8*(1), 27–48. https://doi.org/10.1177/105971230000800103

Diggins, C. A. (2021). Behaviors associated with vocal communication of squirrels. *Ecosphere*, *12*(6), e03572.

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J. J., Ding, M., & Poeppel, D. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, *27*(9), 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002

Dingle, C., Halfwerk, W., & Slabbekoorn, H. (2008). Habitat-dependent song divergence at subspecies level in the grey-breasted wood-wren. *Journal of Evolutionary Biology*, *21*(4), 1079–1089. https://doi.org/10.1111/j.1420-9101.2008.01536.x

Doh, Y., Delfour, F., Augier, E., Glotin, H., Graff, C., & Adam, O. (2018). Bottlenose dolphin (Tursiops truncatus) sonar slacks off before touching a non-alimentary target. *Behavioural Processes*, *157*, 337–345. https://doi.org/10.1016/j.beproc.2018.07.015

Douglas, L. A., Dawson, S. M., & Jaquet, N. (2005). Click rates and silences of sperm whales at Kaikoura, New Zealand. *The Journal of the Acoustical Society of America*, *118*(1), 523–529. https://doi.org/10.1121/1.1937283

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, *22*(1), 567–631.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*(3), 251–288. https://doi.org/10.1016/s0010-0277(00)00106-2

Dray, S., & Dufour, A.-B. (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1–20.

Drouot, V., Goold, J. C., & Gannier, A. (2004). Regional diversity in the social vocalizations of sperm whale in the Mediterranean Sea. *Revue D'Ecologie (La Terre Et La Vie)*, *59*, 545–558.

DuBois, A. L., Nowicki, S., & Searcy, W. A. (2009). Swamp sparrows modulate vocal performance in an aggressive context. *Biology Letters*, *5*(2), 163–165.

Egnor, S. E. R., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences*, *27*(11), 649–654.

Eguiguren, A., Pirotta, E., Cantor, M., Rendell, L., & Whitehead, H. (2019). Habitat use of culturally distinct Galápagos sperm whale Physeter macrocephalus clans. *Marine Ecology Progress Series*, *609*, 257–270. https://doi.org/10.3354/meps12822

Elie, J. E., & Theunissen, F. E. (2016). The vocal repertoire of the domesticated zebra finch: a data-driven approach to decipher the information-bearing acoustic features of communication signals. *Animal Cognition*, *19*(2), 285–315.

Esch, H. C., Sayigh, L. S., & Wells, R. S. (2009). Quantifying parameters of bottlenose dolphin signature whistles. *Marine Mammal Science*, *25*(4), 976–986. https://doi.org/10.1111/j.1748-7692.2009.00289.x

Evans, K., Love, T., & Thurston, S. W. (2015). Outlier identification in model-based cluster analysis. *Journal of Classification*, *32*(1), 63–84. https://doi.org/10.1007/s00357-015-9171-5

Fernandes, M. G. de M. de C. (2016). *Acoustic analysis of slow click function and foraging in sperm whales (Physeter macrocephalus) off Kaikoura, New Zealand*. University of Canterbury.

Filatova, O A, Guzeev, M. A., Fedutin, I. D., Burdin, A. M., & Hoyt, E. (2013). Dependence of killer whale (Orcinus orca) acoustic signals on the type of activity and social context. *Biology Bulletin*, *40*(9), 790–796.

Filatova, Olga A., Deecke, V. B., Ford, J. K. B., Matkin, C. O., Barrett-Lennard, L. G., Guzeev, M. A., Burdin, A. M., & Hoyt, E. (2012). Call diversity in the North Pacific killer whale populations: Implications for dialect evolution and population history. *Animal Behaviour*, *83*(3), 595–603. https://doi.org/10.1016/j.anbehav.2011.12.013

Filatova, Olga A., Ford, J. K. B., Matkin, C. O., Barrett-Lennard, L. G., Burdin, A. M., & Hoyt, E. (2012). Ultrasonic whistles of killer whales (Orcinus orca) recorded in the North Pacific (L). *The Journal of the Acoustical Society of America*, *132*(6), 3618–3621. https://doi.org/10.1121/1.4764874

Fitch, W. T. (2012). The biology and evolution of rhythm: Unravelling a paradox. In P. Rebuschat, M. Rohrmeier, J. A. Hawkins, & I. Cross (Eds.), *Language and Music as Cognitive Systems* (pp. 73–95). Oxford University Press. https://doi.org/https://doi.org/10.1093/acprof:oso/9780199553426.003.0009

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience*, *7*, 1–16. https://doi.org/10.3389/fnsys.2013.00068

Fitch, W. T. (2015). Four principles of bio-musicology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664). https://doi.org/10.1098/rstb.2014.0091

Fonseca, P. J. (2014). Cicada acoustic communication. In *Insect hearing and acoustic communication* (pp. 101–121). Springer.

Foote, A. D., Osborne, R. W., & Hoelzel, A. R. (2004). Whale-call response to masking boat noise. *Nature*, *428*(6986), 910.

Fournet, M. E. H., Gabriele, C. M., Culp, D. C., Sharpe, F., Mellinger, D. K., & Klinck, H. (2018). Some things never change: multi-decadal stability in humpback whale calling repertoire on Southeast Alaskan foraging grounds. *Scientific Reports*, *8*(1), 1–13. https://doi.org/10.1038/s41598-018-31527-x

Fox, A. (2021). How the brainless slime mold stores memories. *Scientific American*. https://www.smithsonianmag.com/smart-news/how-brainless-slime-mold-stores-memories-180977097/

Frantzis, A., & Alexiadou, P. (2008). Male sperm whale (Physeter macrocephalus) coda production and coda-type usage depend on the presence of conspecifics and the behavioural context. *Canadian Journal of Zoology*, *86*(1), 62–75. https://doi.org/10.1139/Z07-114

Freeberg, T. M. (2006). Social complexity can drive vocal complexity: group size influences vocal information in Carolina chickadees. *Psychological Science*, *17*(7), 557–561.

Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1597), 1785–1801. https://doi.org/10.1098/rstb.2011.0213

Frisch, K. von. (2013). *The dance language and orientation of bees*. Harvard University Press. https://doi.org/10.4159/harvard.9780674418776

Gall, G. E. C., & Manser, M. B. (2017). Group cohesion in foraging meerkats: Follow the moving 'vocal hot spot.' *Royal Society Open Science*, *4*(4). https://doi.org/10.1098/rsos.170004

Garland, E. C., Goldizen, A. W., Lilley, M. S., Rekdahl, M. L., Garrigue, C., Constantine, R., Hauser, N. D., Poole, M. M., Robbins, J., & Noad, M. J. (2015). Population structure of humpback whales in the western and central South Pacific Ocean as determined by vocal exchange among populations. *Conservation Biology*, *29*(4), 1198–1207. https://doi.org/https://doi.org/10.1111/cobi.12492

Garland, E. C., Goldizen, A. W., Rekdahl, M. L., Constantine, R., Garrigue, C., Hauser, N. D., Poole, M. M., Robbins, J., & Noad, M. J. (2011). Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Current Biology*, *21*(8), 687–691. https://doi.org/10.1016/j.cub.2011.03.019

Garland, E. C., & McGregor, P. K. (2020). Cultural transmission, evolution, and revolution in vocal displays: Insights from bird and whale song. *Frontiers in Psychology*, *11*, 2387. https://doi.org/10.3389/fpsyg.2020.544929

Gedamke, J. (2004). *Minke whale song, spacing, and acoustic communication on the Great Barrier Reef, Australia*. University of California Santa Cruz.

Gedamke, J., Costa, D. P., & Dunstan, A. (2001). Localization and visual verification of a complex minke whale vocalization. *The Journal of the Acoustical Society of America*, *109*(6), 3038–3047. https://doi.org/10.1121/1.1371763

Geisler, J. H., Colbert, M. W., & Carew, J. L. (2014). A new fossil species supports an early origin for toothed whale echolocation. *Nature*, *508*(7496), 383–386. https://doi.org/10.1038/nature13086

Gerhard, D. (2003). Silence as a cue to rhythm in the analysis of speech and song. *Canadian Acoustics*, *31*(3), 22–23.

Gero, S. (2012). *On the dynamics of social relationships and vocal communication between individuals and social units of sperm whales*. Dalhousie University.

Gero, S., Bøttcher, A., Whitehead, H., & Madsen, P. T. (2016). Socially segregated, sympatric sperm whale clans in the Atlantic Ocean. *Royal Society Open Science*, *3*(6), 160061. https://doi.org/10.1098/rsos.160061

Gero, S., Engelhaupt, D., Rendell, L., & Whitehead, H. (2009). Who cares? Between-group variation in alloparental caregiving in sperm whales. *Behavioral Ecology*, *20*(4), 838–843.

Gero, S., Milligan, M., Rinaldi, C., Francis, P., Gordon, J., Carlson, C., Steffen, A., Tyack, P., Evans, P., & Whitehead, H. (2014). Behavior and social structure of the sperm whales of Dominica, West Indies. *Marine Mammal Science*, *30*(3), 905–922. https://doi.org/10.1111/mms.12086

Gero, S., Whitehead, H., & Rendell, L. (2016). Individual, unit and vocal clan level identity cues in sperm whale codas. *Royal Society Open Science*, *3*, 150372. https://doi.org/10.1098/rsos.150372

Ghazanfar, A. A. (2013). Multisensory vocal communication in primates and the evolution of rhythmic speech. *Behavioral Ecology and Sociobiology*, *67*(9), 1441–1448. https://doi.org/10.1007/s00265-013-1491-z

Gillespie, D. (1997). An acoustic survey for sperm whales in the Southern Ocean Sanctuary conducted from the RSV Aurora Australis. *Reports of the International Whaling Commission*, *47*, 897–907.

Godard, C., Clark, R., Kerr, I., Teglberg Madsen, P., & Payne, R. (2003). Preliminary report on the sperm whale data collected during the Voyage of the Odyssey. *IWC*.

Goodale, E., & Podos, J. (2010). Persistence of song types in Darwin's finches, Geospiza fortis, over four decades. *Biology Letters*, *6*(5), 589–592.

Gordon, J., Leaper, R., Hartley, F. G., & Chappell, O. (1992). Effects of whale-watching vessels on the surface and underwater acoustic behaviour of sperm whales off Kaikoura, New Zealand. *Science and Research Series*, *52*, 1–64.

Gray, P. M., Krause, B., Atema, J., Payne, R., Krumhansl, C., & Baptista, L. (2001). The music of nature and the nature of music. *Science*, *291*(5501), 52–54. https://doi.org/10.1126/science.10.1126/science.1056960

Green, S. R., Mercado, E., Pack, A. A., & Herman, L. M. (2011). Recurring patterns in the songs of humpback whales (Megaptera novaeangliae). *Behavioural Processes*, *86*(2), 284–294. https://doi.org/10.1016/j.beproc.2010.12.014

Gridley, T. (2010). *Geographic and species variation in bottlenose dolphin (Tursiops spp.) signature whistle types*. University of St. Andrews.

Guerra, M., Dawson, S., Sabadel, A., Slooten, E., Somerford, T., Williams, R., Wing, L., & Rayment, W. (2020). Changes in habitat use by a deep-diving predator in response to a coastal earthquake. *Deep-Sea Research Part I: Oceanographic Research Papers*, *January*, 103226. https://doi.org/10.1016/j.dsr.2020.103226

Guerra, M., Hickmott, L., van der Hoop, J., Rayment, W., Leunissen, E., Slooten, E., & Moore, M. (2017). Diverse foraging strategies by a marine top predator: Sperm whales exploit pelagic and demersal habitats in the Kaikoura submarine canyon. *Deep-Sea Research Part I: Oceanographic Research Papers*, *128*, 98–108. https://doi.org/10.1016/j.dsr.2017.08.012

Guinee, L. N., & Payne, K. B. (1988). Rhyme-like repetitions in songs of humpback whales. *Ethology*, *79*(4), 295–306. https://doi.org/10.1111/j.1439-0310.1988.tb00718.x

Haddock, S. H. D., Moline, M. A., & Case, J. F. (2010). Bioluminescence in the sea. *Annual Review of Marine Science*, *2*, 443–493.

Halfwerk, W., Dingle, C., Brinkhuizen, D. M., Poelstra, J. W., Komdeur, J., & Slabbekoorn, H. (2016). Sharp acoustic boundaries across an altitudinal avian hybrid zone despite asymmetric introgression. *Journal of Evolutionary Biology*, *29*(7), 1356–1367. https://doi.org/10.1111/jeb.12876

Handel, S., Todd, S. K., & Zoidis, A. M. (2009). Rhythmic structure in humpback whale (Megaptera novaeangliae) songs: Preliminary implications for song production and perception. *The Journal of the Acoustical Society of America*, *125*(6), EL225–EL230. https://doi.org/10.1121/1.3124712

Handel, S., Todd, S. K., & Zoidis, A. M. (2012). Hierarchical and rhythmic organization in the songs of humpback whales (Megaptera novaeangliae). *Bioacoustics*, *21*(2), 141–156. https://doi.org/10.1080/09524622.2012.668324

Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 543. https://doi.org/10.1037/a0027225

Hansen, M., Wahlberg, M., & Madsen, P. T. (2008). Low-frequency components in harbor porpoise (Phocoena phocoena) clicks: communication signal, by-products, or artifacts? *The Journal of the Acoustical Society of America*, *124*(6), 4059–4068.

Hartbauer, M., & Roemer, H. (2016). Rhythm generation and rhythm perception in insects: the evolution of synchronous choruses. *Frontiers in Neuroscience*, *10*, 223.

Heard, N. A., & Rubin-Delanchy, P. (2018). Choosing between methods of combining p-values. *Biometrika*, *105*(1), 239–246. https://doi.org/10.1093/biomet/asx076

Hersh, T. A., Gero, S., Rendell, L., & Whitehead, H. (2021). Using identity calls to detect structure in acoustic datasets. *Methods in Ecology and Evolution*, *2021*(March), 1–11. https://doi.org/10.1111/2041-210x.13644

Hertenstein, M. J., Verkamp, J. M., Kerestes, A. M., & Holmes, R. M. (2006). The communicative functions of touch in humans, nonhuman primates, and rats: a review and synthesis of the empirical research. *Genetic, Social, and General Psychology Monographs*, *132*(1), 5–94.

Herzing, D. (2015). Synchronous and rhythmic vocalizations and correlated underwater behavior of free-ranging Atlantic spotted dolphins (Stenella frontalis) and bottlenose dolphins (Tursiops truncatus) in the Bahamas. *Animal Behavior and Cognition*, *2*(1), 14–29. https://doi.org/10.12966/abc.02.02.2015

Hickey, P., Merseal, H., Patel, A. D., & Race, E. (2020). Memory in time: Neural tracking of low-frequency rhythm dynamically modulates memory formation. *NeuroImage*, *213*, 116693. https://doi.org/10.1016/j.neuroimage.2020.116693

Hijmans, R. J., Williams, E., & Vennes, C. (2015). *Geosphere: spherical trigonometry*.

Hinskens, F. (2014). Despite or because of intensive contact? Internal, external and extralinguistic aspects of divergence in modern dialects and ethnolects of Dutch. In K. Braunmuller, S. Hoder, & K. Kuhl (Eds.), *Stability and Divergence in Language Contact: Factors and Mechanisms* (Vol. 16, pp. 109–140).

Hinskens, F., Auer, P., & Kerswill, P. (2005). The study of dialect convergence and divergence: conceptual and methodological considerations. *Dialect Change: Convergence and Divergence in European Languages*, 1–48.

Honkola, T., Ruokolainen, K., Syrjänen, K. J. J., Leino, U. P., Tammi, I., Wahlberg, N., & Vesakoski, O. (2018). Evolution within a language: Environmental differences contribute to divergence of dialect groups. *BMC Evolutionary Biology*, *18*(1), 1–15. https://doi.org/10.1186/s12862-018-1238-6

Hooker, S. K., & Whitehead, H. (2002). Click characteristics of northern bottlenose whales (Hyperoodon ampullatus). *Marine Mammal Science*, *18*(1), 69–80. https://doi.org/10.1111/j.1748-7692.2002.tb01019.x

Hopkins, C D. (2009). Electrical perception and communication. *Encyclopedia of Neuroscience*, *3*, 813–831.

Hopkins, Carl D. (1988). Neuroethology of electric communication. *Annual Review of Neuroscience*, *11*(1), 497–535.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, *27*(6), 949–960. https://doi.org/10.1521/soco.2009.27.6.949

Huijser, L. A. E., Estrade, V., Webster, I., Mouysset, L., Cadinouche, A., & Dulau-Drouot, V. (2020). Vocal repertoires and insights into social structure of sperm whales (Physeter macrocephalus) in Mauritius, southwestern Indian Ocean. *Marine Mammal Science*, *36*(2), 638–657. https://doi.org/10.1111/mms.12673

Huron, D. (2008). *Sweet anticipation: Music and the psychology of expectation*. MIT Press. https://doi.org/10.7551/mitpress/6575.001.0001

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*(2), 297–307. https://doi.org/https://doi.org/10.1093/biomet/76.2.297

Hyland Bruno, J. (2017). *Song rhythm development in zebra finches*. City University of New York.

I'Anson Price, R., & Grüter, C. (2015). Why, when and where did honey bee dance communication evolve? *Frontiers in Ecology and Evolution*, *3*, 125.

Ihara, Y. (2011). Evolution of culture-dependent discriminate sociality: a gene–culture coevolutionary model. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1566), 889–900.

Ioup, G. E., Ioup, J. W., Pflug, L. A., Tashmukhambetov, A. M., Sidorovskaia, N. A., Schexnayder, P., Tiemann, C. O., Bernstein, A., Kuczaj, S. A., Rayborn, G. H., Newcomb, J. J., Carlson, R., & Ekimov, A. (2009). EARS buoy applications by LADC: I. Marine animal acoustics. *OCEANS 2009*, 1–9. https://doi.org/https://doi.org/10.23919/oceans.2009.5422190

Irvine, L., Palacios, D. M., Urbán, J., & Mate, B. (2017). Sperm whale dive behavior characteristics derived from intermediate-duration archival tag data. *Ecology and Evolution*, *7*(19), 7822–7837.

Irwin, E. (2012). The spermaceti candle and the american whaling industry. *Historia*, *21*, 45–53.

Jabr, F. (2012). How brainless slime molds redefine intelligence. *Nature*. https://doi.org/10.1038/nature.2012.11811

Jacoby, N., & McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*, *27*(3), 359–370. https://doi.org/10.1016/j.cub.2016.12.031

Janes, S. W., & Ryker, L. (2013). Rapid change in a type I song dialect of Hermit Warblers (Setophaga occidentalis). *The Auk*, *130*(1), 30–35.

Jang, Y., & Gerhardt, H. C. (2006). Divergence in female calling song discrimination between sympatric and allopatric populations of the southern wood cricket Gryllus fultoni (Orthoptera: Gryllidae). *Behavioral Ecology and Sociobiology*, *60*(2), 150–158.

Janik, V. M., Sayigh, L. S., & Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences*, *103*(21), 8293–8297. https://doi.org/10.1073/pnas.0509918103

Janik, Vincent M. (2009). Whale song. *Current Biology*, *19*(3), R109–R111. https://doi.org/10.1016/j.cub.2008.11.026

Janik, Vincent M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, *28*, 60–65. https://doi.org/10.1016/j.conb.2014.06.010

Jaquet, N. (1996). How spatial and temporal scales influence understanding of sperm whale distribution: a review. *Mammal Review*, *26*(1), 51–65.

Jaquet, N., Dawson, S., & Douglas, L. (2001). Vocal behavior of male sperm whales: Why do they click? *The Journal of the Acoustical Society of America*, *109*(5), 2254–2259. https://doi.org/10.1121/1.1360718

Johndro, H., Jacobs, L., Patel, A. D., & Race, E. (2019). Temporal predictions provided by musical rhythm influence visual memory encoding. *Acta Psychologica*, *200*, 102923. https://doi.org/10.1016/j.actpsy.2019.102923

Johnson, M., Madsen, P. T., Zimmer, W. M. X., Aguilar De Soto, N., & Tyack, P. L. (2006). Foraging Blainville's beaked whales (Mesoplodon densirostris) produce distinct click types matched to different phases of echolocation. *Journal of Experimental Biology*, *209*(24), 5038–5050. https://doi.org/10.1242/jeb.02596

Jones, A., & Ward, E. V. (2019). Rhythmic temporal structure at encoding enhances recognition memory. *Journal of Cognitive Neuroscience*, *31*(10), 1549–1562. https://doi.org/10.1162/jocn

Jones, Mari R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(5), 1059–1073. https://doi.org/10.1037/0096-1523.7.5.1059

Jones, Mari Riess, Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, *13*(4), 313–319. https://doi.org/10.1111/1467-9280.00458

Jones, Mari Riess, & Pfordresher, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology*, *51*(4), 271–290. https://doi.org/10.1037/1196-1961.51.4.271

Keith, S. A., & Bull, J. W. (2017). Animal culture impacts species' capacity to realise climate-driven range shifts. *Ecography*, *40*(2), 296–304.

Keller, E. F. (2010). *The mirage of a space between nature and nurture*. Duke University Press.

Kello, C. T., Dalla Bella, S., Mede, B., & Balasubramaniam, R. (2017). Hierarchical temporal structure in music, speech and animal vocalizations: Jazz is like a conversation, humpbacks sing like hermit thrushes. *Journal of the Royal Society Interface*, *14*(135), 20170231. https://doi.org/10.1098/rsif.2017.0231

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., Bohn, K., Cao, Y., Carter, G., Cäsar, C., Coen, M., Deruiter, S. L., Doyle, L., Edelman, S., Ferrer-i-Cancho, R., Freeberg, T. M., Garland, E. C., Gustison, M., Harley, H. E., … Zamora-Gutierrez, V. (2016). Acoustic sequences in non-human animals: A tutorial review and prospectus. *Biological Reviews*, *91*(1), 13–52. https://doi.org/10.1111/brv.12160

Kershenbaum, A., Root-Gutteridge, H., Habib, B., Koler-Matznick, J., Mitchell, B., Palacios, V., & Waller, S. (2016). Disentangling canid howls across multiple species and subspecies: Structure in a complex communication channel. *Behavioural Processes*, *124*, 149–157. https://doi.org/10.1016/j.beproc.2016.01.006

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364. https://doi.org/10.1016/j.evolhumbehav.2010.04.004

Knörnschild, M. (2014). Vocal production learning in bats. *Current Opinion in Neurobiology*, *28*, 80–85.

Knörnschild, M., Fernandez, A. A., & Nagy, M. (2020). Vocal information and the navigation of social decisions in bats: Is social complexity linked to vocal complexity? *Functional Ecology*, *34*(2), 322–331.

Kobayashi, H., Whitehead, H., & Amano, M. (2020). Long-term associations among male sperm whales (Physeter macrocephalus). *Plos One*, *15*(12), e0244204.

Konrad, C. M., Frasier, T. R., Rendell, L., Whitehead, H., & Gero, S. (2018). Kinship and association do not explain vocal repertoire variation among individual sperm whales or social units. *Animal Behaviour*, *145*, 131–140.

Konrad, C. M., Frasier, T. R., Whitehead, H., & Gero, S. (2019). Kin selection and allocare in sperm whales. *Behavioral Ecology*, *30*(1), 194–201.

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences*, *22*(10), 896–910. https://doi.org/10.1016/j.tics.2018.08.002

Krams, I., Krama, T., Freeberg, T. M., Kullberg, C., & Lucas, J. R. (2012). Linking social complexity and vocal complexity: a parid perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1597), 1879–1891.

Langus, A., Mehler, J., & Nespor, M. (2017). Rhythm in language acquisition. *Neuroscience & Biobehavioral Reviews*, *81*, 158–166. https://doi.org/10.1016/j.neubiorev.2016.12.012

Laplanche, C., Adam, O., Lopatka, M., & Motsch, J.-F. (2005). Male sperm whale acoustic behavior observed from multipaths at a single hydrophone. *The Journal of the Acoustical Society of America*, *118*(4), 2677–2687. https://doi.org/10.1121/1.2033567

Large, E. W., & Jones, M. R. (1999). They dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159. https://doi.org/10.1037/0033-295x.106.1.119

Lattenkamp, E. Z., & Vernes, S. C. (2018). Vocal learning: A language-relevant trait in need of a broad cross-species approach. *Current Opinion in Behavioral Sciences*, *21*, 209–215. https://doi.org/10.1016/j.cobeha.2018.04.007

Lavery, T. J., Roudnew, B., Gill, P., Seymour, J., Seuront, L., Johnson, G., Mitchell, J. G., & Smetacek, V. (2010). Iron defecation by sperm whales stimulates carbon export in the Southern Ocean. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1699), 3527–3531.

Le Bot, O., Mars, J. I., Gervaise, C., & Simard, Y. (2015). Rhythmic analysis for click train detection and source separation with examples on beluga whales. *Applied Acoustics*, *95*, 37–49. https://doi.org/10.1016/j.apacoust.2015.02.005

Le Bot, Olivier, Bonnel, J., Mars, J. I., & Gervaise, C. (2013). Odontocete click train deinterleaving using a single hydrophone and rhythm analysis. *Proceedings of Meetings on Acoustics*, *19*(1). https://doi.org/10.1121/1.4800655

LeDuc, R. G., Robertson, K. M., & Pitman, R. L. (2008). Mitochondrial sequence divergence among Antarctic killer whale ecotypes is consistent with multiple species. *Biology Letters*, *4*(4), 426–429. https://doi.org/10.1098/rsbl.2008.0168

Lee, J.-H., Podos, J., & Sung, H.-C. (2019). Distinct patterns of geographic variation for different song components in Daurian Redstarts Phoenicurus auroreus. *Bird Study*, *66*(1), 73–82.

Leow, L.-A., & Grahn, J. A. (2014). Neural mechanisms of rhythm perception: present findings and future directions. *Neurobiology of Interval Timing*, 325–338.

Li, K., Torres, C. E., Thomas, K., Rossi, L. F., & Shen, C.-C. (2011). Slime mold inspired routing protocols for wireless sensor networks. *Swarm Intelligence*, *5*(3), 183–223.

Lieberman, P. H., Klatt, D. H., & Wilson, W. H. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science*, *164*(3884), 1185–1187.

Linnenschmidt, M., Teilmann, J., Akamatsu, T., Dietz, R., & Miller, L. A. (2013). Biosonar, dive, and foraging activity of satellite tracked harbor porpoises (Phocoena phocoena). *Marine Mammal Science*, *29*(2), 77–97. https://doi.org/10.1111/j.1748-7692.2012.00592.x

Lipkind, D., Geambasu, A., & Levelt, C. C. (2020). The development of structured vocalizations in songbirds and humans: a comparative analysis. *Topics in Cognitive Science*, *12*(3), 894–909.

Loffredo, C. A., & Borgia, G. (1986). Male courtship vocalizations as cues for mate choice in the satin bowerbird (Ptilonorhynchus violaceus). *The Auk*, *103*(1), 189–195.

Lycett, S. J. (2019). Confirmation of the role of geographic isolation by distance in among-tribe variations in beadwork designs and manufacture on the High Plains. *Archaeological and Anthropological Sciences*, *11*(6), 2837–2847.

Lyrholm, T., Leimar, O., Johanneson, B., & Gyllensten, U. (1999). Sex–biased dispersal in sperm whales: Contrasting mitochondrial and nuclear genetic structure of global populations. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1417), 347–354. https://doi.org/https://doi.org/10.1098/rspb.1999.0644

Mann, N. I., Dingess, K. A., & Slater, P. J. B. (2006). Antiphonal four-part synchronized chorusing in a Neotropical wren. *Biology Letters*, *2*(1), 1–4. https://doi.org/10.1098/rsbl.2005.0373

Marcoux, M., Rendell, L., & Whitehead, H. (2007). Indications of fitness differences among vocal clans of sperm whales. *Behavioral Ecology and Sociobiology*, *61*(7), 1093–1098. https://doi.org/https://doi.org/10.1007/s00265-006-0342-6

Marcoux, M., Whitehead, H., & Rendell, L. (2007). Sperm whale feeding variation by location, year, social group and clan: Evidence from stable isotopes. *Marine Ecology Progress Series*, *333*, 309–314. https://doi.org/10.3354/meps333309

Margoliash, D. (1983). Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *Journal of Neuroscience*, *3*(5), 1039–1057. https://doi.org/10.1523/jneurosci.03-05-01039.1983

Mauch, M., MacCallum, R. M., Levy, M., & Leroi, A. M. (2015). The evolution of popular music: USA 1960–2010. *Royal Society Open Science*, *2*(5), 150081.

McAuley, J. D. (2010). Tempo and rhythm. In Mari Riess Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (Vol. 36, pp. 165–199). Springer Science & Business Media. https://doi.org/10.1007/978-1-4419-6114-3_6

McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biology Letters*, *1*(4), 381–385.

McDonald, M. A., Hildebrand, J. A., & Mesnick, S. (2009). Worldwide decline in tonal frequencies of blue whale songs. *Endangered Species Research*, *9*(1), 13–21. https://doi.org/10.3354/esr00217

McDonald, M. A., Mesnick, S. L., & Hildebrand, J. A. (2006). Biogeographic characterisation of blue whale song worldwide: Using song to identify populations. *Journal of Cetacean Research and Management*, *8*(1), 55–65. https://doi.org/10.1029/2006WR005124.DOI

McElreath, R., Boyd, R., & Richerson, P. J. (2003). Shared norms and the evolution of ethnic markers. *Current Anthropology*, *44*(1), 122–130.

McGowen, M. R., Gatesy, J., & Wildman, D. E. (2014). Molecular evolution tracks macroevolutionary transitions in Cetacea. *Trends in Ecology & Evolution*, *29*(6), 336–346.

McNeil, W. H. (1997). *Keeping together in time: Dance and drill in human history*. Harvard University Press. https://doi.org/10.2307/j.ctvjf9wq6

Mellinger, D. K., & Clark, C. W. (2003). Blue whale (Balaenoptera musculus) sounds from the North Atlantic. *The Journal of the Acoustical Society of America*, *114*(2), 1108–1119. https://doi.org/10.1121/1.1593066

Merker, B. H., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, *45*(1), 4–17. https://doi.org/10.1016/j.cortex.2008.06.011

Miller, B. S., Growcott, A., Slooten, E., & Dawson, S. M. (2013). Acoustically derived growth rates of sperm whales (Physeter macrocephalus) in Kaikoura, New Zealand. *The Journal of the Acoustical Society of America*, *134*(3), 2438–2445.

Miller, P. J. O., Shapiro, A. D., Tyack, P. L., & Solow, A. R. (2004). Call-type matching in vocal exchanges of free-ranging resident killer whales, Orcinus orca. *Animal Behaviour*, *67*(6), 1099–1107. https://doi.org/10.1016/j.anbehav.2003.06.017

Mizroch, S. A., & Rice, D. W. (2013). Ocean nomads: distribution and movements of sperm whales in the North Pacific shown by whaling data and Discovery marks. *Marine Mammal Science*, *29*(2), E136–E165.

Moffett, M. W. (2013). Human identity and the evolution of societies. *Human Nature*, *24*(3), 219–267. https://doi.org/10.1007/s12110-013-9170-3

Møhl, B., Wahlberg, M., Madsen, P. T., Miller, L. A., & Surlykke, A. (2000). Sperm whale clicks: Directionality and source level revisited. *The Journal of the Acoustical Society of America*, *107*(1), 638–648.

Mooney, T. A., Kaplan, M. B., & Lammers, M. O. (2016). Singing whales generate high levels of particle motion: Implications for acoustic communication and hearing? *Biology Letters*, *12*(11), 20160381. https://doi.org/10.1098/rsbl.2016.0381

Mooney, T. A., Nachtigall, P. E., Taylor, K. A., Rasmussen, M. H., & Miller, L. A. (2009). Auditory temporal resolution of a wild white-beaked dolphin (Lagenorhynchus albirostris). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *195*(4), 375–384. https://doi.org/10.1007/s00359-009-0415-x

Moore, B. L., Connor, R. C., Allen, S. J., Krützen, M., & King, S. L. (2020). Acoustic coordination by allied male dolphins in a cooperative context. *Proceedings of the Royal Society B*, *287*(1924), 20192944. https://doi.org/10.1098/rspb.2019.2944

Moore, K. E., Watkins, W. A., & Tyack, P. L. (1993). Pattern similarity in shared codas from sperm whales (Physeter catodon). *Marine Mammal Science*, *9*(1), 1–9. https://doi.org/https://doi.org/10.1111/j.1748-7692.1993.tb00421.x

Moran, P. A., Hunt, J., Mitchell, C., Ritchie, M. G., & Bailey, N. W. (2020). Sexual selection and population divergence III: Interspecific and intraspecific variation in mating signals. *Journal of Evolutionary Biology*, *33*(7), 990–1005. https://doi.org/10.1111/jeb.13631

Moritz, C. (1994). Defining 'evolutionarily significant units' for conservation. *Trends in Ecology and Evolutionology & Evolution*, *9*(10), 373–375.

Mumtaz, K., & Duraiswamy, K. (2010). A novel density based improved k-means clustering algorithm - Dbkmeans. *International Journal on Computer Science and Engineering*, *02*(02), 213–218.

Mundinger, P. C. (1980). Animal cultures and a general theory of cultural evolution. *Ethology and Sociobiology*, *1*(3), 183–223.

Nachtigall, P. E., & Supin, A. Y. (2015). Conditioned frequency-dependent hearing sensitivity reduction in the bottlenose dolphin (Tursiops truncatus). *Journal of Experimental Biology*, *218*(7), 999–1005. https://doi.org/10.1242/jeb.114066

Nachtigall, P. E., Supin, A. Y., Smith, A. B., & Pacini, A. F. (2016). Expectancy and conditioned hearing levels in the bottlenose dolphin (Tursiops truncatus). *Journal of Experimental Biology*, *219*(6), 844–850. https://doi.org/10.1242/jeb.133777

Nakahara, F., & Miyazaki, N. (2011). Vocal exchanges of signature whistles in bottlenose dolphins (Tursiops truncatus). *Journal of Ethology*, *29*(2), 309–320. https://doi.org/10.1007/s10164-010-0259-4

Nettle, D. (2009). Ecological influences on human behavioural diversity: a review of recent findings. *Trends in Ecology & Evolution*, *24*(11), 618–624.

Nieder, A., & Mooney, R. (2020). The neurobiology of innate, volitional and learned vocalizations in mammals and birds. *Philosophical Transactions of the Royal Society B*, *375*(1789), 20190054.

Norris, T. F., Oswald, J., Yack, T., Ferguson, E., Hom-Weaver, C., Dunleavy, K., Coates, S., & Dominello, T. (2012). *An analysis of acoustic data from the Mariana Islands Sea Turtle and Cetacean Survey (MISTCS)*.

Norton, P., & Scharff, C. (2016). "Bird song metronomics": Isochronous organization of zebra finch song rhythm. *Frontiers in Neuroscience*, *10*, 309. https://doi.org/10.3389/fnins.2016.00309

Oleson, E. M., Barlow, J., Gordon, J., Rankin, S., & Hildebrand, J. A. (2003). Low frequency calls of Bryde's whales. *Marine Mammal Science*, *19*(2), 407–419. https://doi.org/10.1111/j.1748-7692.2003.tb01119.x

Oleson, E. M., Calambokidis, J., Burgess, W. C., McDonald, M. A., LeDuc, C. A., & Hildebrand, J. A. (2007). Behavioral context of call production by eastern North Pacific blue whales. *Marine Ecology Progress Series*, *330*, 269–284. https://doi.org/10.3354/meps330269

Oleson, E. M., Širović, A., Bayless, A. R., & Hildebr, J. A. (2014). Synchronous seasonal change in fin whale song in the North Pacific. *PLoS ONE*, *9*(12), 1–18. https://doi.org/10.1371/journal.pone.0115678

Oliveira, C., Wahlberg, M., Silva, M. A., Johnson, M., Antunes, R., Wisniewska, D. M., Fais, A., Gonçalves, J., & Madsen, P. T. (2016). Sperm whale codas may encode individuality as well as clan identity. *The Journal of the Acoustical Society of America*, *139*(5), 2860–2869. https://doi.org/10.1121/1.4949478

Ord, T. J., & Garcia-Porta, J. (2012). Is sociality required for the evolution of communicative complexity? Evidence weighed against alternative hypotheses in diverse taxonomic groups. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1597), 1811–1828.

Oswald, J. N., Au, W. W. L., & Duennebier, F. (2011). Minke whale (Balaenoptera acutorostrata) boings detected at the Station ALOHA Cabled Observatory. *The Journal of the Acoustical Society of America*, *129*(5), 3353–3360. https://doi.org/10.1121/1.3575555

Oswald, J. N., Walmsley, S. F., Casey, C., Fregosi, S., Southall, B., & Janik, V. M. (2021). Species information in whistle frequency modulation patterns of common dolphins. *Philosophical Transactions of the Royal Society B*, *376*(1836), 20210046.

Pagel, M., Atkinson, Q. D., & Meade, A. (2007). Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature*, *449*(7163), 717–720. https://doi.org/10.1038/nature06176

Pante, E., & Simon-Bouhet, B. (2013). marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS One*, *8*(9), e73051. https://doi.org/10.1371/journal.pone.0073051

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*(4), 409–464. https://doi.org/10.2307/40285633

Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195123753.001.0001

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, *87*(1), B35–B45.

Pavan, G., Hayward, T. J., Borsani, J. F., Priano, M., Manghi, M., Fossati, C., & Gordon, J. (2000). Time patterns of sperm whale codas recorded in the Mediterranean Sea 1985-1996. *The Journal of the Acoustical Society of America*, *107*(6), 3487–3495. https://doi.org/10.1121/1.429419

Payne, R., & McVay, S. (1971). Songs of humpback whales. *Science*, *173*(3997), 585–597. https://doi.org/10.1126/science.173.3997.585

Pebesma, E. J. (2018). Simple features for R: Standardized support for spatial vector data. *R J.*, *10*(1), 439.

Pereira, A., Harris, D., Tyack, P., & Matias, L. (2020). Fin whale acoustic presence and song characteristics in seas to the southwest of Portugal. *The Journal of the Acoustical Society of America*, *147*(4), 2235–2249. https://doi.org/10.1121/10.0001066

Petkov, C. I., & Jarvis, E. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, *4*, 12.

Pfennig, K., & Pfennig, D. (2009). Character displacement: ecological and reproductive responses to a common evolutionary problem. *The Quarterly Review of Biology*, *84*(3), 253–276.

Philbrick, N. (2001). *In the heart of the sea: The tragedy of the whaleship Essex*. Penguin.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(9), 3526–3529. https://doi.org/10.1073/pnas.1012551108

Pirotta, E., Matthiopoulos, J., MacKenzie, M., Scott-Hayward, L., & Rendell, L. (2011). Modelling sperm whale habitat preference: A novel approach combining transect and follow data. *Marine Ecology Progress Series*, *436*, 257–272. https://doi.org/10.3354/meps09236

Polak, R., Jacoby, N., Fischinger, T., Goldberg, D., Holzapfel, A., & London, J. (2018). Rhythmic prototypes across cultures: A comparative study of tapping synchronization. *Music Perception: An Interdisciplinary Journal*, *36*(1), 1–23. https://doi.org/10.1525/mp.2018.36.1.1

Popov, V. V, & Supin, A. Y. (1998). Auditory evoked responses to rhythmic sound pulses in dolphins. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *183*(4), 519–524. https://doi.org/10.1007/s003590050277

Price, J. J., & Lanyon, S. M. (2002). Reconstructing the evolution of complex bird song in the oropendolas. *Evolution*, *56*(7), 1514–1529.

Punzo, A., Mazza, A., & McNicholas, P. D. (2018). ContaminatedMixt: An R package for fitting parsimonious mixtures of multivariate contaminated normal distributions. *Journal of Statistical Software*, *85*(10), 1–25. https://doi.org/10.18637/jss.v085.i10

Punzo, A., & McNicholas, P. D. (2016). Parsimonious mixtures of multivariate contaminated normal distributions. *Biometrical Journal*, *58*(6), 1506–1537. https://doi.org/10.1002/bimj.201500144

Rankin, S., Oswald, J., Barlow, J., & Lammers, M. (2007). Patterned burst-pulse vocalizations of the northern right whale dolphin, Lissodelphis borealis. *The Journal of the Acoustical Society of America*, *121*(2), 1213–1218. https://doi.org/10.1121/1.2404919

Ravignani, Andrea. (2019a). Rhythm and synchrony in animal movement and communication. *Current Zoology*, *65*(1), 77–81. https://doi.org/10.1093/cz/zoy087

Ravignani, Andrea. (2019b). Timing of antisynchronous calling: A case study in a harbor seal pup (Phoca vitulina). *Journal of Comparative Psychology*, *133*(2), 272–277. https://doi.org/10.1037/com0000160

Ravignani, Andrea, Bowling, D., & Fitch, W. T. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology*, *5*, 1–15. https://doi.org/10.3389/fpsyg.2014.01118

Ravignani, Andrea, Dalla Bella, S., Falk, S., Kello, C. T., Noriega, F., & Kotz, S. A. (2019). Rhythm in speech and animal vocalizations: A cross-species perspective. *Annals of the New York Academy of Sciences*, *1453*, 79–98. https://doi.org/10.1111/nyas.14166

Ravignani, Andrea, Fitch, W. T., Hanke, F. D., Heinrich, T., Hurgitsch, B., Kotz, S. A., Scharff, C., Stoeger, A. S., & Boer, B. de. (2016). What pinnipeds have to say about human speech, music, and the evolution of rhythm. *Frontiers in Neuroscience*, *10*, 274. https://doi.org/10.3389/fnins.2016.00274

Ravignani, Andrea, Kello, C. T., De Reus, K., Kotz, S. A., Dalla Bella, S., Mendez-Arostegui, M., Rapado-Tamarit, B., Rubio-Garcia, A., & De Boer, B. (2019). Ontogeny of vocal rhythms in harbor seal pups: An exploratory study. *Current Zoology*, *65*(1), 107–120. https://doi.org/10.1093/cz/zoy055

Ravignani, Andrea, & Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, *8*, 1820. https://doi.org/10.3389/fpsyg.2017.01820

Ravignani, Andrea, & Norton, P. (2017). Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. *Journal of Language Evolution*, *2*(1), 4–19. https://doi.org/10.1093/jole/lzx002

Ravignani, Andrea, Thompson, B., Grossi, T., Delgado, T., & Kirby, S. (2018). Evolving building blocks of rhythm: How human cognition creates music via cultural transmission. *Annals of the New York Academy of Sciences*, *1423*(1), 176–187. https://doi.org/10.1111/nyas.13610

Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, *65*(3), 519–530.

Recalde-Salas, A., Salgado Kent, C. P., Parsons, M. J. G., Marley, S. A., & McCauley, R. D. (2014). Non-song vocalizations of pygmy blue whales in Geographe Bay, Western Australia. *The Journal of the Acoustical Society of America*, *135*(5), EL213–EL218. https://doi.org/10.1121/1.4871581

Reidenberg, J. S., & Laitman, J. T. (2018). Anatomy of underwater sound production with a focus on ultrasonic vocalization in toothed whales including dolphins and porpoises. In S. M. Brudzynski (Ed.), *Handbook of Behavioral Neuroscience* (Vol. 25, pp. 509–519). Elsevier. https://doi.org/10.1016/b978-0-12-809600-0.00047-0

Rekdahl, M. L., Dunlop, R. A., Goldizen, A. W., Garland, E. C., Biassoni, N., Miller, P., & Noad, M. J. (2015). Non-song social call bouts of migrating humpback whales. *The Journal of the Acoustical Society of America*, *137*(6), 3042–3053. https://doi.org/10.1121/1.4921280

Rekdahl, M. L., Dunlop, R. A., Noad, M. J., & Goldizen, A. W. (2013). Temporal stability and change in the social call repertoire of migrating humpback whales. *The Journal of the Acoustical Society of America*, *133*(3), 1785–1795. https://doi.org/10.1121/1.4789941

Remeslo, A., Yukhov, V., Bolstad, K., & Laptikhovsky, V. (2019). Distribution and biology of the colossal squid, Mesonychoteuthis hamiltoni: New data from depredation in toothfish fisheries and sperm whale stomach contents. *Deep Sea Research Part I: Oceanographic Research Papers*, *147*, 121–127.

Rendell, L. E., & Whitehead, H. (2003a). Comparing repertoires of sperm whale codas: A multiple methods approach. *Bioacoustics*, *14*(1), 61–81. https://doi.org/10.1080/09524622.2003.9753513

Rendell, L. E., & Whitehead, H. (2003b). Vocal clans in sperm whales (Physeter macrocephalus). *Proceedings of the Royal Society of London Series B: Biological Sciences*, *270*(1512), 225–231. https://doi.org/10.1098/rspb.2002.2239

Rendell, L., & Frantzis, A. (2016). Mediterranean sperm whales, Physeter macrocephalus: The precarious state of a lost tribe. *Advances in Marine Biology*, *75*, 37–74. https://doi.org/10.1016/bs.amb.2016.08.001

Rendell, L., Simião, S., Brotons, J. M., Airoldi, S., Fasano, D., & Gannier, A. (2014). Abundance and movements of sperm whales in the western Mediterranean basin. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *24*(S1), 31–40. https://doi.org/10.1002/aqc.2426

Rendell, L, Mesnick, S. L., Dalebout, M. L., Burtenshaw, J., & Whitehead, H. (2012). Can genetic differences explain vocal dialect variation in sperm whales, Physeter macrocephalus? *Behavior Genetics*, *42*(2), 332–343. https://doi.org/10.1007/s10519-011-9513-y

Rendell, L., & Whitehead, H. (2005). Spatial and temporal variation in sperm whale coda vocalizations: Stable usage and local dialects. *Animal Behaviour*, *70*(1), 191–198. https://doi.org/10.1016/j.anbehav.2005.03.001

Rendell, L., Whitehead, H., & Coakes, A. (2005). Do breeding male sperm whales show preferences among vocal clans of females? *Marine Mammal Science*, *21*(2), 317–322.

Rhinelander, M. Q., & Dawson, S. M. (2004). Measuring sperm whales from their clicks: Stability of interpulse intervals and validation that they indicate whale length. *The Journal of the Acoustical Society of America*, *115*(4), 1826–1831.

Ribeiro Jr., P. J., & Diggle, P. J. (2001). The geoR package. *R News*, *1*(2), 14–18.

Rice, D. W. (1989). Sperm whale. Physeter macrocephalus Linnaeus, 1758. In S. H. Ridgway & R. J. Harrison (Eds.), *Handbook of Marine Mammals* (Vol. 4, pp. 177–233). Academic Press.

Ridgway, S. H. (2011). Neural time and movement time in choice of whistle or pulse burst responses to different auditory stimuli by dolphins. *The Journal of the Acoustical Society of America*, *129*(2), 1073–1080. https://doi.org/10.1121/1.3523431

Riera, A., Rountree, R. A., Mouy, X., Ford, J. K., & Juanes, F. (2016). Effects of anthropogenic noise on fishes at the SGaan Kinghlas-Bowie Seamount Marine Protected Area. *Proceedings of Meetings on Acoustics 4ENAL*, *27*(1), 10005.

Riesch, R., Barrett-Lennard, L. G., Ellis, G. M., Ford, J. K. B., & Deecke, V. B. (2012). Cultural traditions and the evolution of reproductive isolation: Ecological speciation in killer whales? *Biological Journal of the Linnean Society*, *106*(1), 1–17. https://doi.org/10.1111/j.1095-8312.2012.01872.x

Riesch, R., Ford, J. K. B., & Thomsen, F. (2006). Stability and group specificity of stereotyped whistles in resident killer whales, Orcinus orca, off British Columbia. *Animal Behaviour*, *71*(1), 79–91. https://doi.org/10.1016/j.anbehav.2005.03.026

Rimmele, J., Jolsvai, H., & Sussman, E. (2011). Auditory target detection is affected by implicit temporal and spatial expectations. *Journal of Cognitive Neuroscience*, *23*(5), 1136–1147. https://doi.org/10.1162/jocn.2010.21437

Roberts, G. (2008). Language and the free-rider problem: An experimental paradigm. *Biological Theory*, *3*(2), 174–183.

Roeske, T. C., Tchernichovski, O., Poeppel, D., & Jacoby, N. (2020). Categorical rhythms are shared between songbirds and humans. *Current Biology*, *30*(18), 3544–3555. https://doi.org/10.1016/j.cub.2020.06.072

Roper, C. F. E., & Boss, K. J. (1982). The giant squid. *Scientific American*, *246*(4), 96–105.

Rubenstein, C. (1963). 22—The lubrication of spun yarns. *Journal of the Textile Institute Transactions*, *54*(6), T234–T253. https://doi.org/10.1080/19447026308660176

Ryan, S. J. (2006). The role of culture in conservation planning for small or endangered populations. *Conservation Biology*, *20*(4), 1321–1324.

Ryder, O. A. (1986). Species conservation and systematics: the dilemma of subspecies. *Trends Ecol. Evol.*, *1*, 9–10.

Sainburg, T., Thielk, M., & Gentner, T. Q. (2020). Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires. *PLoS Computational Biology*, *16*(10), e1008228.

Sakai, M., Morisaka, T., Kogi, K., Hishii, T., & Kohshima, S. (2010). Fine-scale analysis of synchronous breathing in wild Indo-Pacific bottlenose dolphins (Tursiops aduncus). *Behavioural Processes*, *83*(1), 48–53. https://doi.org/10.1016/j.beproc.2009.10.001

Sayigh, L., Daher, M. A., Allen, J., Gordon, H., Joyce, K., Stuhlmann, C., & Tyack, P. (2016). The Watkins Marine Mammal Sound Database: An online, freely accessible resource. *Proceedings of Meetings on Acoustics*, 040013. https://doi.org/10.1121/2.0000358

Sayigh, L., Quick, N., Hastie, G., & Tyack, P. (2013). Repeated call types in short-finned pilot whales, Globicephala macrorhynchus. *Marine Mammal Science*, *29*(2), 312–324. https://doi.org/10.1111/j.1748-7692.2012.00577.x

Sayigh, L. S., Tyack, P. L., Wells, R. S., & Scott, M. D. (1990). Signature whistles of free-ranging bottlenose dolphins Tursiops truncatus: stability and mother-offspring comparisons. *Behavioral Ecology and Sociobiology*, *26*(4), 247–260.

Schehka, S., Esser, K.-H., & Zimmermann, E. (2007). Acoustical expression of arousal in conflict situations in tree shrews (Tupaia belangeri). *Journal of Comparative Physiology A*, *193*(8), 845–852.

Schneider, J. N., & Mercado III, E. (2019). Characterizing the rhythm and tempo of sound production by singing whales. *Bioacoustics*, *28*(3), 239–256. https://doi.org/10.1080/09524622.2018.1428827

Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2008). Overlapping and matching of codas in vocal interactions between sperm whales: Insights into communication function. *Animal Behaviour*, *76*(6), 1977–1988. https://doi.org/10.1016/j.anbehav.2008.07.032

Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2011). Individual vocal production in a sperm whale (Physeter macrocephalus) social unit. *Marine Mammal Science*, *27*(1), 149–166. https://doi.org/10.1111/j.1748-7692.2010.00399.x

Schuppli, C., & van Schaik, C. P. (2019). Animal cultures: How we've only seen the tip of the iceberg. *Evolutionary Human Sciences*, *1*, 1–13. https://doi.org/10.1017/ehs.2019.1

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464. https://doi.org/https://doi.org/10.1214/aos/1176344136

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, *8*(1), 289.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Senigaglia, V., de Stephanis, R., Verborgh, P., & Lusseau, D. (2012). The role of synchronized swimming as affiliative and anti-predatory behavior in long-finned pilot whales. *Behavioural Processes*, *91*(1), 8–14. https://doi.org/10.1016/j.beproc.2012.04.011

Serrano, A., & Terhune, J. (2002). Stability of the underwater vocal repertoire of harp seals (Pagophilus groenlandicus). *Aquatic Mammals*, *28*(1), 93–101.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Animal Behaviour*, *28*(4), 1070–1094.

Sharpe, F. A. (2001). *Social foraging of the southeast Alaskan humpback whale, Megaptera novaeangliae*. Simon Fraser University.

Shireman, E., Steinley, D., & Brusco, M. J. (2017). Examining the effect of initialization strategies on the performance of Gaussian mixture modeling. *Behavior Research Methods*, *49*(1), 282–293. https://doi.org/https://doi.org/10.3758/s13428-015-0697-6

Simard, S. W. (2018). Mycorrhizal networks facilitate tree communication, learning, and memory. In *Memory and learning in plants* (pp. 191–213). Springer.

Simonis, A. E., Baumann-Pickering, S., Oleson, E., Melcón, M. L., Gassmann, M., Wiggins, S. M., & Hildebrand, J. A. (2012). High-frequency modulated signals of killer whales (Orcinus orca) in the North Pacific. *The Journal of the Acoustical Society of America*, *131*(4), EL295–EL301. https://doi.org/10.1121/1.3690963

Širović, A., Hildebrand, J. A., & Wiggins, S. M. (2007). Blue and fin whale call source levels and propagation range in the Southern Ocean. *The Journal of the Acoustical Society of America*, *122*(2), 1208–1215. https://doi.org/10.1121/1.2749452

Širović, A., Oleson, E. M., Buccowich, J., Rice, A., & Bayless, A. R. (2017). Fin whale song variability in southern California and the Gulf of California. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-09979-4

Smith, W. J. (1991). Singing is based on two markedly different kinds of signaling. *Journal of Theoretical Biology*, *152*(2), 241–253. https://doi.org/10.1016/S0022-5193(05)80455-2

Spierings, M. J., & ten Cate, C. (2016). Zebra finches as a model species to understand the roots of rhythm. *Frontiers in Neuroscience*, *10*, 345. https://doi.org/10.3389/fnins.2016.00345

Stadler, K. (2018). *cultevo: Tools, measures and statistical tests for cultural evolution* (1.0.2). https://kevinstadler.github.io/cultevo/

Stafford, K. M., Moore, S. E., Laidre, K. L., & Heide-Jørgensen, M. P. (2008). Bowhead whale springtime song off West Greenland. *The Journal of the Acoustical Society of America*, *124*(5), 3315–3323. https://doi.org/10.1121/1.2980443

Stafford, K. M., Nieukirk, S. L., & Fox, C. G. (2001). Geographic and seasonal variation of blue whale calls in the North Pacific. *Journal of Cetacean Research and Management*, *3*(1), 65–76.

Steiner, L., Lamoni, L., Plata, M. A., Jensen, S.-K., Lettevall, E., & Gordon, J. (2012). A link between male sperm whales, Physeter macrocephalus, of the Azores and Norway. *Journal of the Marine Biological Association of the United Kingdom*, *92*(8), 1751–1756.

Straley, J. M., Schorr, G. S., Thode, A. M., Calambokidis, J., Lunsford, C. R., Chenoweth, E. M., Connell, V. M. O., & Andrews, R. D. (2014). Depredating sperm whales in the Gulf of Alaska: local habitat use and long distance movements across putative population boundaries. *Endangered Species Research*, *24*(2), 125–135.

Suthers, R. A., Thomas, S. P., & Suthers, B. J. (1972). Respiration, wing-beat and ultrasonic pulse emission in an echo-locating bat. *Journal of Experimental Biology*, *56*(1), 37–48. https://doi.org/10.1242/jeb.56.1.37

Szymanski, M. D., Supin, A. Y., Bain, D. E., & Henry, K. R. (1998). Killer whale (Orcinus orca) auditory evoked potentials to rhythmic clicks. *Marine Mammal Science*, *14*(4), 676–691. https://doi.org/10.1111/j.1748-7692.1998.tb00756.x

Tarr, B., Launay, J., Cohen, E., & Dunbar, R. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, *11*(10), 20150767. https://doi.org/10.1098/rsbl.2015.0767

Taylor, B., Baird, R., Barlow, J., Dawson, S. M., Ford, J., Mead, J. G., & Pitman, R. L. (2019). Physeter macrocephalus (amended version of 2008 assessment). *The IUCN Red List of Threatened Species*, e.T41755A160983555. https://doi.org/10.2305/IUCN.UK.2008.RLTS.T41755A160983555.en

ten Cate, C., Spierings, M., Hubert, J., & Honing, H. (2016). Can birds perceive rhythmic patterns? A review and experiments on a songbird and a parrot species. *Frontiers in Psychology*, *7*(MAY), 1–14. https://doi.org/10.3389/fpsyg.2016.00730

Terhune, J. M. (1994). Geographical variation of harp seal underwater vocalizations. *Canadian Journal of Zoology*, *72*(5), 892–897. https://doi.org/10.1139/z94-121

Thavabalasingam, S., O'Neil, E. B., Zeng, Z., & Lee, A. C. H. (2016). Recognition memory is improved by a structured temporal framework during encoding. *Frontiers in Psychology*, *6*, 2062. https://doi.org/10.3389/fpsyg.2015.02062

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift Für Tierpsychologie*, *20*(4), 410–433. https://doi.org/10.1111/j.1439-0310.1963.tb01161.x

Tønnesen, P., Gero, S., Ladegaard, M., Johnson, M., & Madsen, P. T. (2018). First-year sperm whale calves echolocate and perform long, deep dives. *Behavioral Ecology and Sociobiology*, *72*(10), 1–15. https://doi.org/10.1007/s00265-018-2570-y

Tønnesen, P., Oliveira, C., Johnson, M., & Madsen, P. T. (2020). The long-range echo scene of the sperm whale biosonar. *Biology Letters*, *16*(8), 20200134. https://doi.org/10.1098/rsbl.2020.0134

Treese, S. A. (2018). Historical luminous intensity. In *History and Measurement of the Base and Derived Units* (pp. 927–945). Springer. https://doi.org/10.1007/978-3-319-77577-7_13

Tremblay, C. J., Van Parijs, S. M., & Cholewiak, D. (2019). 50 to 30-Hz triplet and singlet down sweep vocalizations produced by sei whales (Balaenoptera borealis) in the western North Atlantic Ocean. *The Journal of the Acoustical Society of America*, *145*(6), 3351–3358. https://doi.org/10.1121/1.5110713

Tyack, P L. (1998). Acoustic communication under the sea. In *Animal acoustic communication* (pp. 163–220). Springer. https://doi.org/10.1007/978-3-642-76220-8

Tyack, Peter L, & Sayigh, L. S. (1997). Vocal learning in cetaceans. In C. T. Snowdon & M. Hausberger (Eds.), *Social Influences on Vocal Development* (pp. 208–233). Cambridge University Press. https://doi.org/10.1017/cbo9780511758843.011

United Nations Environmental Program. (2014). Conservation implications of cetacean culture. Convention on Migratory Species Resolution. *UNEP/CMS/ COP11/Doc. 23.2.4*. http://www.cms.int/sites/default/files/document/COP11_Doc_23_2_4_Conservation_Implications_Cetacean_En.pdf

Vachon, F., Hersh, T. A., Rendell, L., Gero, S., & Whitehead, H. (n.d.). Ocean nomads or island specialists? Eastern Caribbean sperm whales show culturally driven fine-scale habitat partitioning which contrasts with the scale of Pacific populations. *Submitted*.

Van Cise, A. M., Mahaffy, S. D., Baird, R. W., Mooney, T. A., & Barlow, J. (2018). Song of my people: dialect differences among sympatric social groups of short-finned pilot whales in Hawai'i. *Behavioral Ecology and Sociobiology*, *72*(12), 193. https://doi.org/10.1007/s00265-018-2596-1

Van Parijs, S. M., & Corkeron, P. J. (2001). Vocalizations and behaviour of Pacific humpback dolphins Sousa chinensis. *Ethology*, *107*(8), 701–716. https://doi.org/10.1046/j.1439-0310.2001.00714.x

Vargas, M. F. N. R. (2017). *Revealing structure in vocalisations of parrots and social whales*. Georg-August-Universität Göttingen.

Varughese, J. C., Moser, D., Thenius, R., Wotawa, F., & Schmickl, T. (2019). swarmfstaxis: Borrowing a swarm communication mechanism from fireflies and slime mold. In *Complex Adaptive Systems* (pp. 213–222). Springer. https://doi.org/10.1007/978-3-030-20309-2_10

Wadewitz, P., Hammerschmidt, K., Battaglia, D., Witt, A., Wolf, F., & Fischer, J. (2015). Characterizing vocal repertoires—Hard vs. soft classification approaches. *PloS One*, *10*(4), e0125785. https://doi.org/10.1371/journal.pone.0125785

Waller, K. M., Thorson, P. H., Sawyer, K. A., Wilkinson, D. L., Fulling, G. L., Norris, T. F., Yack, T., Hall, C., Fertl, D., Gehring, P., Knight, K., Rexstad, E., See, J., & Watterson, J. C. (2007). *Marine mammal and sea turtle survey and density estimates for Guam and the Commonwealth of the Northern Mariana Islands*.

Waters, C. M., & Bassler, B. L. (2005). Quorum sensing: Cell-to-cell communication in bacteria. *Annual Review of Cell and Developmental Biology*, *21*, 319–346. https://doi.org/10.1146/annurev.cellbio.21.012704.131001

Watkins, W. A. (1985). Investigations of sperm whale acoustic behaviors in the southeast Caribbean. *Cetology*, *49*, 1–15.

Watkins, W. A., & Moore, K. E. (1982). An underwater acoustic survey for sperm whales (Physeter catodon) and other cetaceans in the southeast Caribbean. *Cetology*, *46*, 1–7.

Watkins, W. A., & Schevill, W. E. (1977). Sperm whale codas. *Journal of the Acoustical Society of America*, *62*(6), 1485–1490. https://doi.org/10.1121/1.381678

Weilgart, L., & Whitehead, H. (1997). Group-specific dialects and geographic variation in coda repertoire in South Pacific sperm whales. *Behavioral Ecology and Sociobiology*, *40*(5), 277–285. https://doi.org/10.1007/s002650050343

Weirathmueller, M. J., Stafford, K. M., Wilcock, W. S. D., Hilmo, R. S., Dziak, R. P., & Tréhu, A. M. (2017). Spatial and temporal trends in fin whale vocalizations recorded in the NE Pacific Ocean between 2003-2013. *PLoS ONE*, *12*(10), 1–24. https://doi.org/10.1371/journal.pone.0186127

Wey, T., Blumstein, D. T., Shen, W., & Jordán, F. (2008). Social network analysis of animal behaviour: a promising tool for the study of sociality. *Animal Behaviour*, *75*(2), 333–344. https://doi.org/10.1016/j.anbehav.2007.06.020

Whitehead, H. (1996). Babysitting, dive synchrony, and indications of alloparental care in sperm whales. *Behavioral Ecology and Sociobiology*, *38*(4), 237–244. https://doi.org/10.1007/s002650050238

Whitehead, H. (2001). Analysis of animal movement using opportunistic individual identifications: Application to sperm whales. *Ecology*, *82*(5), 1417–1432. https://doi.org/10.2307/2679999

Whitehead, H. (2002). Estimates of the current global population size and historical trajectory for sperm whales. *Marine Ecology Progress Series*, *242*, 295–304. https://doi.org/10.3354/meps242295

Whitehead, H. (2003). *Sperm whales: social evolution in the ocean*. University of Chicago Press.

Whitehead, H. (2018). Sperm whale: Physeter macrocephalus. In B. Wursig, J. G. M. Thewissen, & K. M. Kovacs (Eds.), *Encyclopedia of Marine Mammals* (pp. 919–925). Elsevier. https://doi.org/10.1016/C2015-0-00820-6

Whitehead, H., Antunes, R., Gero, S., Wong, S. N. P., Engelhaupt, D., & Rendell, L. (2012). Multilevel societies of female sperm whales (Physeter macrocephalus) in the Atlantic and Pacific: Why are they so different? *International Journal of Primatology*, *33*(5), 1142–1164. https://doi.org/10.1007/s10764-012-9598-z

Whitehead, H., Coakes, A., Jaquet, N., & Lusseau, S. (2008). Movements of sperm whales in the tropical Pacific. *Marine Ecology Progress Series*, *361*, 291–300. https://doi.org/10.3354/meps07412

Whitehead, H., & Hersh, T. A. (n.d.). Posterior probabilities of membership in acoustic identity clades. *Submitted*.

Whitehead, H., & Rendell, L. (2004). Movements, habitat use and feeding success of cultural clans of South Pacific sperm whales. *Journal of Animal Ecology*, *73*(1), 190–196. https://doi.org/10.1111/j.1365-2656.2004.00798.x

Whitehead, H., & Rendell, L. (2014). *The cultural lives of whales and dolphins*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226187426.001.0001

Whitehead, H., Rendell, L., Osborne, R. W., & Würsig, B. (2004). Culture and conservation of non-humans with reference to whales and dolphins: review and new directions. *Biological Conservation*, *120*(3), 427–437. https://doi.org/10.1016/j.biocon.2004.03.017

Whitehead, H., Smith, T. D., & Rendell, L. (2021). Adaptation of sperm whales to open-boat whalers: Rapid social learning on a large scale? *Biology Letters*, *17*(3). https://doi.org/10.1098/rsbl.2021.0030

Whitehead, H., Waters, S., & Lyrholm, T. (1991). Social organization of female sperm whales and their offspring: constant companions and casual acquaintances. *Behavioral Ecology and Sociobiology*, *29*(5), 385–389. https://doi.org/10.1007/BF00165964

Whitehead, H., & Weilgart, L. (2000). The sperm whale: social females and roving males. *Cetacean Societies: Field Studies of Dolphins and Whales*, 154–172.

Whiten, A. (2017). A second inheritance system: the extension of biology through culture. *Interface Focus*, *7*(5), 20160142. https://doi.org/10.1098/rsfs.2016.0142

Whiten, A. (2021). The burgeoning reach of animal culture. *Science*, *372*(6537). https://doi.org/10.1126/science.abe6514

Widdess, R. (2013). Schemas and improvisation in Indian music. In R. Kempson, C. Howes, & M. Orwin (Eds.), *Language, Music, and Interaction* (pp. 197–209). College Publications.

Wieland, M., Jones, A., & Renn, S. C. P. (2010). Changing durations of southern resident killer whale (Orcinus orca) discrete calls between two periods spanning 28 years. *Marine Mammal Science*, *26*(1), 195–201. https://doi.org/10.1111/j.1748-7692.2009.00351.x

Wiley, R. H. (2013). Signal detection, noise, and the evolution of communication. In H. Brumm (Ed.), *Animal Communication and Noise* (2nd ed., pp. 7–31). Springer. https://doi.org/10.1007/978-3-642-41494-7

Wilkins, M. R., Seddon, N., & Safran, R. J. (2013). Evolutionary divergence in acoustic signals: Causes and consequences. *Trends in Ecology & Evolution*, *28*(3), 156–166. https://doi.org/https://doi.org/10.1016/j.tree.2012.10.002

Wilkinson, G. S. (2003). Social and vocal complexity in bats. In *Animal social complexity: Intelligence, culture, and individualized societies* (pp. 322–341). Harvard University Press Cambridge, MA. https://doi.org/10.4159/harvard.9780674419131.c24

Williams, H., Levin, I. I., Norris, D. R., Newman, A. E. M., & Wheelwright, N. T. (2013). Three decades of cultural evolution in Savannah sparrow songs. *Animal Behaviour*, *85*(1), 213–223. https://doi.org/10.1016/j.anbehav.2012.10.028

Williams, H., & Staples, K. (1992). Syllable chunking in zebra finch (Taeniopygia guttata) song. *Journal of Comparative Psychology*, *106*(3), 278. https://doi.org/10.1037/0735-7036.106.3.278

Wilson, M., & Cook, P. F. (2016). Rhythmic entrainment: Why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychonomic Bulletin and Review*, *23*(6), 1647–1659. https://doi.org/10.3758/s13423-016-1013-x

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

Yamamoto, Y., Akamatsu, T., da Silva, V. M. F., & Kohshima, S. (2016). Local habitat use by botos (Amazon river dolphins, Inia geoffrensis) using passive acoustic methods. *Marine Mammal Science*, *32*(1), 220–240. https://doi.org/10.1111/mms.12252

Yurk, H., Barrett-Lennard, L., Ford, J. K. B., & Matkin, C. O. (2002). Cultural transmission within maternal lineages: vocal clans in resident killer whales in southern Alaska. *Animal Behaviour*, *63*(6), 1103–1119. https://doi.org/10.1006/anbe.2002.3012

Zahavi, A., & Zahavi, A. (1999). *The handicap principle: A missing piece of Darwin's puzzle*. Oxford University Press.

Zaugg, S., van der Schaar, M., Houégnigan, L., & André, M. (2013). Extraction of pulse repetition intervals from sperm whale click trains for ocean acoustic data mining. *The Journal of the Acoustical Society of America*, *133*(2), 902–911. https://doi.org/10.1121/1.4773278

Zimmer, W. M. X., Johnson, M. P., Madsen, P. T., & Tyack, P. L. (2005). Echolocation clicks of free-ranging Cuvier's beaked whales (Ziphius cavirostris). *The Journal of the Acoustical Society of America*, *117*(6), 3919–3927. https://doi.org/10.1121/1.1910225

Zwamborn, E. M. J., & Whitehead, H. (2017). Repeated call sequences and behavioural context in long-finned pilot whales off Cape Breton, Nova Scotia, Canada. *Bioacoustics*, *26*(2), 169–183. https://doi.org/10.1080/09524622.2016.1233457

# APPENDIX A – CHAPTER 2 SUPPLEMENT

**A2.1:** *Method S2.1* – Literature review methods

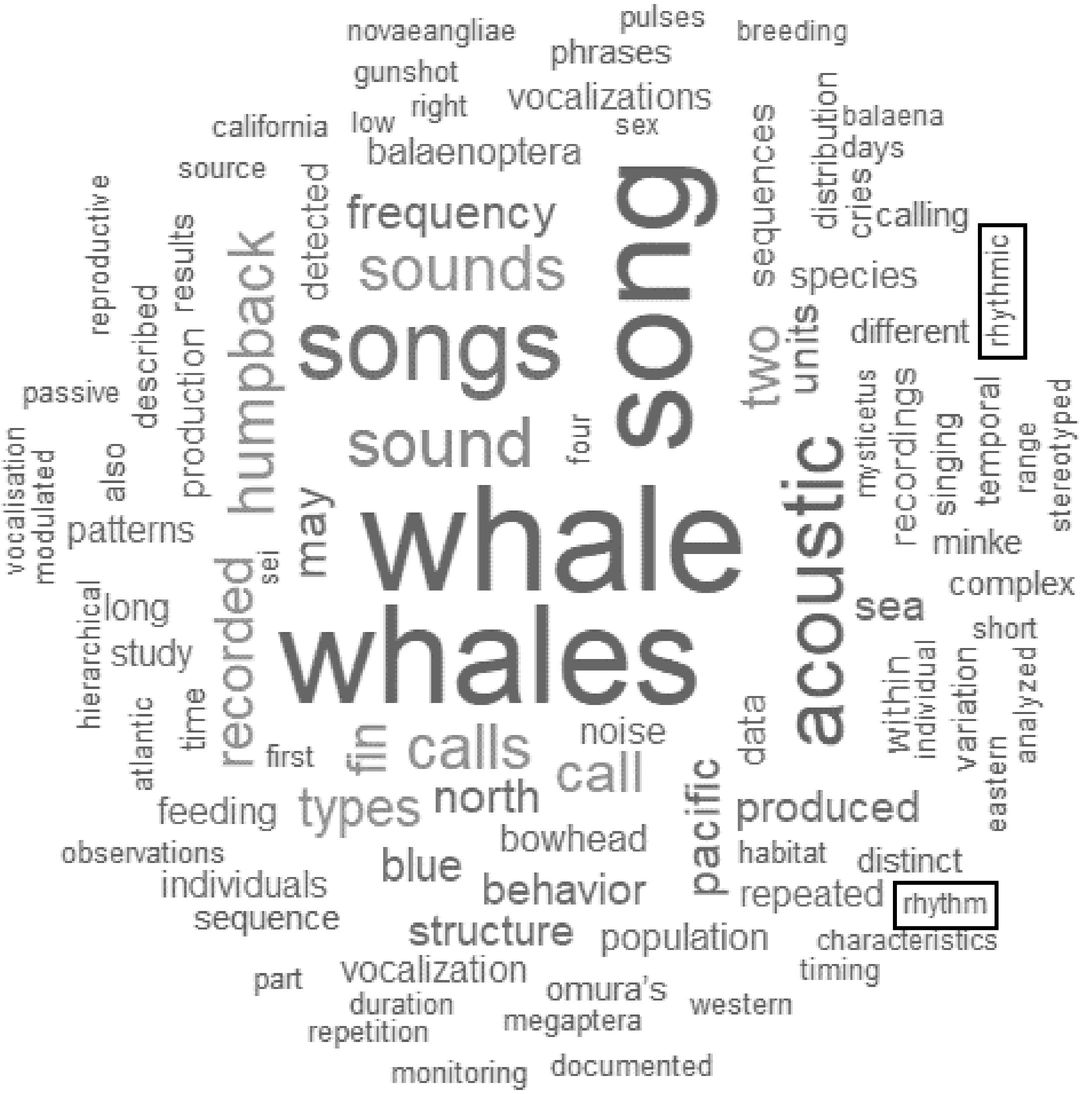
An initial scoping literature review on rhythm in marine mammal vocalizations was conducted using a keyword search (Table S2.3) of four abstract and citation databases: Biological Abstracts, Web of Science Core Collection, ProQuest, and Scopus. 1,552 studies were imported into the review management software Covidence. 550 studies were automatically identified as duplicates and removed, leaving 1,002 studies for title and abstract screening. During screening, 107 studies were identified as relevant and subjected to a full text review. Of those, 72 were excluded based on established criteria.

Many of the exclusions (n=23) were pinniped studies; the scope of our review initially included pinnipeds but was later restricted to cetaceans. Twenty studies were about rhythmic processing in the odontocete brain, one was about rhythm in humans, and one was about the cross-species approach to rhythm research. Seventeen of the exclusions touched on cetacean vocalizations or acoustics more generally but did not explicitly quantify rhythm. The remaining ten excluded studies were either conference abstracts with no associated publication (n=7) or newspaper articles (n=3).

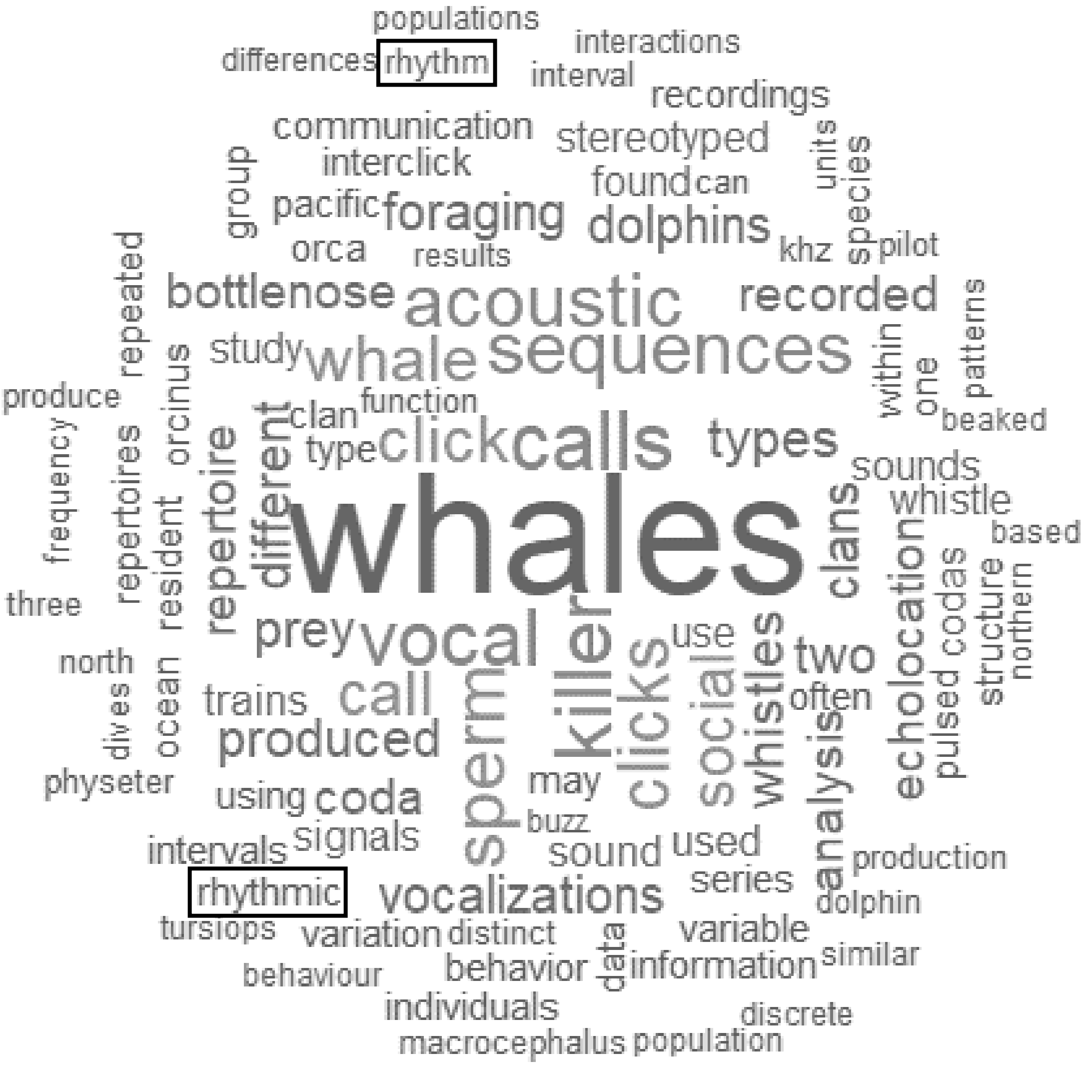
Of the 35 studies that remained after full text review, 24 were about rhythm in odontocete vocalizations and 11 were about rhythm in mysticete vocalizations. Relevant data from the 35 studies were extracted and incorporated into Tables 2.2/S2.1 (mysticete vocal rhythms), Tables 2.3/S2.2 (odontocete vocal rhythms), or into the text of this paper.

It quickly became clear that our keyword search strategy was missing many relevant publications on cetacean vocal rhythm, often because publications that assessed temporal aspects of vocalizations did so without invoking the word ‘rhythm’ (see Figures S2.1/S2.2). For that reason, we also reviewed relevant references cited in the 107 studies that underwent full text review and did extensive “hand searches” for cetacean vocal rhythm research using Google Scholar. Given the primary objective of this review—to provide an overview of the general presence of and trends in cetacean vocal rhythm—and coupled with a lack of uniformity in how cetacean researchers describe rhythmic research, a strictly systematic literature review was not feasible.

**A2.2:** *Figure S2.1* – Mysticete publication word cloud. This cloud shows the 100 most frequently used words in the titles, keywords, and abstracts of the 26 papers from which mysticete rhythmic information was extracted (Tables 2.2/S2.1). Each word was made anywhere from 6 to 83 times, with font size corresponding to usage. Boxes denote words related to rhythm. “Rhythmic” and “rhythm” occurred 7 and 6 times, respectively, while “rhythmically” and “rhythmical” each occurred only once and are not featured in the word cloud. This figure was produced in R version 3.6.1 using the ‘wordcloud’ package.



**A2.3:** *Figure S2.2* – Odontocete publication word cloud. This cloud shows the 100 most frequently used words in the titles, keywords, and abstracts of the 39 papers from which odontocete rhythmic information was extracted (Tables 2.3/S2.2). Each word was made anywhere from 10 to 121 times, with font size corresponding to usage. Boxes denote words related to rhythm. "Rhythmic" and "rhythm" occurred 13 and 11 times, respectively, while "rhythms" occurred two times and is not featured in the word cloud. This figure was produced in R version 3.6.1 using the 'wordcloud' package.



**A2.4:** *Table S2.1* – Detailed examples of definitive and preliminary isochronous and heterochronous rhythm in mysticete vocalizations. For each vocalization, 'Context' refers to the known or hypothesized behavioral context in which that vocalization is produced. 'Unit(s)' refers to the acoustic unit of interest, and rhythm is considered at the level of the inter-unit interval. For both isochrony and heterochrony, the 'Evidence?' column is populated with either **Y**es (evidence was found in the studies examined) or **N**o (evidence was not found in the studies examined). The 'Type(s) of evidence' are **Q**uantitative, **D**escriptive, **V**isual (e.g. spectrograms, chronograms, etc.), and **U**npublished. For the most convincing evidence, references are provided in the 'Citations' column and 'Specific example(s)' are given. Examples can include figures, tables, and interval information (presented as mean±SD with the coefficient of variation in parentheses unless otherwise specified). For examples of heterochronous rhythm, interval ratios were calculated from the average interval durations when possible. This supplemental Excel file can be accessed via Dalspace or via the Open Science Framework (https://osf.io/rb2ea/) and is titled:

HershThesis_AppendixA2.4_TableS2.1_MysticeteVocalRhythm.xlsx

**A2.5:** *Table S2.2* – Detailed examples of definitive and preliminary isochronous and heterochronous rhythm in odontocete vocalizations. For each vocalization, 'Context' refers to the known or hypothesized behavioral context in which that vocalization is produced. 'Unit(s)' refers to the acoustic unit of interest, and rhythm is considered at the level of the inter-unit interval. For both isochrony and heterochrony, the 'Evidence?' column is populated with either **Y**es (evidence was found in the studies examined) or **N**o (evidence was not found in the studies examined). The 'Type(s) of evidence' are **Q**uantitative, **D**escriptive, **V**isual (e.g. spectrograms, chronograms, etc.), and **U**npublished. For the most convincing evidence, references are provided in the 'Citations' column and 'Specific example(s)' are given. Examples can include figures, tables, and interval information (presented as mean±SD with the coefficient of variation in parentheses unless otherwise specified). For examples of heterochronous rhythm, interval ratios were calculated from the average interval durations when possible. This supplemental Excel file can be accessed via Dalspace or via the Open Science Framework (https://osf.io/rb2ea/) and is titled:

HershThesis_AppendixA2.5_TableS2.2_OdontoceteVocalRhythm

**A2.6:** *Table S2.3* – Keywords used to search four abstract and citation databases for research on rhythm in marine mammal vocalizations.

| **Database(s)** | **Keywords** |
|---|---|
| Biological Abstracts, Web of Science Core Collection | (cetacea* OR dolphin* OR whal* OR porpois* OR pinni* OR “seal” OR “seals” OR "sea lion*" OR mysticet* OR odontocet* OR "marine mamm*" OR otariid* OR odoben* OR phocid*) AND (rhythm*) NOT (melatonin) |
| ProQuest | ab(cetacea* OR dolphin* OR whal* OR porpois* OR pinni* OR "seal" OR "seals" OR "sea lion*" OR mysticet* OR odontocet* OR "marine mamm*" OR otariid* OR odoben* OR phocid*) AND ab(rhythm*) NOT ab(melatonin) |
| Scopus | TITLE-ABS-KEY ( cetacea* OR dolphin* OR whal* OR porpois* OR pinni* OR "seal" OR "seals" OR "sea lion*" OR mysticet* OR odontocet* OR "marine mamm*" OR otariid* OR odoben* OR phocid* ) AND TITLE-ABS-KEY ( rhythm* ) AND NOT TITLE-ABS-KEY ( melatonin ) |

# APPENDIX B – CHAPTER 3 SUPPLEMENT

**B3.1:** *Method S3.1* – Extracting sperm whale codas from the Watkins Marine Mammal Sound Database

The Watkins Marine Mammal Sound Database (Sayigh et al., 2016) has 308 sperm whale master tapes. One is labelled as a Galápagos Islands recording and 53 are labelled as eastern Caribbean recordings. The Galápagos recording and a majority of the eastern Caribbean recordings cannot be automatically opened due to errors that occurred when the recordings were converted from the original .KAY file format to .wav files (personal communication, Dr. Laela Sayigh, Woods Hole Oceanographic Institution). However, the files can still be opened by importing them into Audacity (version 2.3.0) as raw data and specifying the file sampling rate. The digitization parameters for each master tape .KAY file, including sampling rate, are binarily encoded in the file header. We accessed these headers using a custom MATLAB (version R2020a) script (accessible at https://osf.io/5fter/), which allowed us to correctly specify the sampling rate for each file and save it as a new, uncorrupted .wav file. All files were then audited for codas. Coda parameters (including the number of clicks and inter-click intervals; ICIs) were extracted using a custom software, called Coda Sorter, which was written by Kristian Beedholm (Marine Bioacoustics Lab, Aarhus University). Coda Sorter is implemented in LabView and run in MATLAB.

**B3.2:** *Method S3.2* – Descriptions and processing of test datasets

*Sperm whales*

Due to inconsistent past marking of extremely short and long codas, and because our goal was to identify clan-specific, frequently used identity codas, only codas with three to nine clicks were analyzed. Codas with more than nine clicks made up, on average, only 2.40±4.21% of each location dataset included in this study. In regions where longer codas are common (e.g. Brazil, Amorim et al., 2020), this range can be increased.

All codas recorded from a single photo-identified group of sperm whales in a year were compiled into a single repertoire for all of the Pacific locations, the Balearic Islands, the Gulf of Mexico, and the Atlantic coast of Panama (for photo-identification methods and definitions, see Cantor et al., 2016; Pirotta et al., 2011; Rendell et al., 2014; Rendell & Whitehead, 2003). For the eastern Caribbean data from Dominica, where the long-term sperm whale social structure is well-documented (Gero et al., 2014), all codas recorded from a single known social unit of sperm whales within a year were compiled into one repertoire. If the identity of recorded whales was unknown or multiple groups/units were present, all of the codas recorded on a given day were combined into a single repertoire to reduce autocorrelation and minimize the possibility of well-sampled groups/units being designated as clans due to group- or unit-specific codas (Antunes et al., 2011; Gero, Whitehead, et al., 2016; Oliveira et al., 2016). Only repertoires with at least 25 codas were analyzed. For each sperm whale repertoire, it is unknown which individuals within the group produced each coda or how many individuals were vocalizing at any given time. Examples of sperm whale coda spectrograms can be viewed in Figures 1, 2, and 3 of Watkins & Schevill (1977).

*Grey-breasted wood-wrens*

Field and data collection methods are detailed in the original study (Halfwerk et al., 2016). The dataset contains acoustic metrics for 471 averaged song types from 66 male wrens belonging to two subspecies (*Henicorhina leucophrys hilaris* and *Henicorhina leucophrys leucophrys)*. Based on genotyping, 22 males were parental *H. l. hilaris*, 29 were parental *H. l. leucophrys*, six were first-generation (F1) hybrids, and nine were second-generation males that were backcrossed between an F1 hybrid and either *H. l. hilaris* (n=4) or *H. l. leucophrys* (n=2). Each song type has measurements for 22 temporal and spectral acoustic characteristics. Previous work found that averaged note peak frequency ($F_{\text{peak}}$), minimum song frequency ($F_{\text{min}}$), and maximum song frequency ($F_{\text{max}}$) differed significantly between male *H. l. hilaris* and *H. l. leucophrys* song (Dingle et al., 2008). For each song type, we created a frequency vector of $F_{\text{peak}}$, $F_{\text{min}}$, and $F_{\text{max}}$ and logged the vectors prior to clustering to account for differences in scale of the frequency measures. Only males that had at least five song types were included in our analysis, which reduced the dataset to 396 song types and 41 males (*H. l. hilaris*=8, *H. l. leucophrys*=22, F1=5, *H. l. hilaris*/F1 backcross=5, *H. l. leucophrys*/F1 backcross=1). Each repertoire thus contains the songs from a single individual. Examples of song spectrograms can be viewed in Figure 4 of Halfwerk et al. (2016).

*Australian field crickets*

Field and data collection methods are detailed in the original study (Moran et al., 2020). The dataset is comprised of calling song data for male *Teleogryllus* crickets (127 *T. commodus*; 131 *T. oceanicus*). These lab-reared, first-generation crickets were derived from wild-caught individuals from 16 field sites. Each cricket's song was recorded five times and 13 song traits were extracted. The 258 songs in the dataset represent the mean of each cricket's five songs. For each song, we created an interval vector that included four of the 13 song traits: chirp pulse length, chirp interpulse interval, chirp-trill interval, and trill pulse length. We chose these traits because they were measured in the same units (seconds), had similar ranges of variation, and looked like promising candidates for species discrimination (see Figure S3 in Moran et al., 2020). Because we only had access to one averaged set of song measurements per male, we used parental field site as the repertoire grouping variable instead of individual. Each repertoire thus contains the songs from multiple individuals. Despite species sympatry at four of the 16 sites, all first-generation males derived from sympatric sites were *T. oceanicus* (Moran et al., 2020). For our purposes, eight sites were thus unambiguously *T. commodus* (sites SV, UQ, CH, BL, MV, CC, BN, AM) and the other eight were unambiguously *T. oceanicus* (KH, DV, JC, PL, YP, RH, TS, HB) (see Moran et al., 2020 for site details). Song schematics showing the various song traits can be viewed in Figure 2 of Moran et al. (2020).

### **B3.3:** *Method S3.3* – Validation of call categories

*Sperm whales*

A variety of approaches, each with their own assumptions and limitations, have been used to classify sperm whale codas into types and delineate repertoires into clans in the past. In the first report of sperm whale vocal clans, Rendell & Whitehead (2003) compared repertoires of codas using multivariate similarity measures and classified codas into categorical types using *k*-means clustering of standardized ICIs. Similar approaches were used to delineate sperm whale clans off Japan (Amano et al., 2014) and to look for individually distinctive features in codas recorded off Dominica (Antunes et al., 2011; Schulz et al., 2011). Due to initial emphasis on coda rhythm rather than tempo (Moore et al., 1993), many sperm whale studies standardized coda ICIs by total coda duration (e.g. Amano et al., 2014; Rendell & Whitehead, 2003; Weilgart & Whitehead, 1997). Additional research suggested that both rhythm and tempo are informative coda features (Antunes et al., 2011; Frantzis & Alexiadou, 2008; Gero, Whitehead, et al., 2016), and more recent work has used absolute coda ICIs (Amorim et al., 2020; Huijser et al., 2020; Oliveira et al., 2016).

*K*-means clustering as a call classification strategy has several drawbacks, a significant one being that *k*—the number of clusters—must be assigned *a priori*. Techniques exist to help with this assignation (e.g. variance ratio criterion; Caliński & Harabasz, 1974) but do not always provide consistent, unambiguous results (Huijser et al., 2020; Rendell & Whitehead, 2003a) and the choice of *k* can be arbitrary or subjective. *K*-means clustering also forces the data into spherical, equal-sized Voronoi cells; is sensitive to initial conditions and outliers; and struggles at higher dimensions (Mumtaz & Duraiswamy, 2010; Oliveira et al., 2016).

More recently, Oliveira et al. (2016) used a combination of principal components analysis and observer classification to divide codas into categorical types. A drawback of this approach is that observer classification of codas can be time consuming, subjective, and inconsistent across observers (Rendell & Whitehead, 2003a). Other studies have used the OPTICS hierarchical clustering algorithm (Ankerst et al., 1999) to divide codas into categorical types (Cantor et al., 2016; Gero, Bøttcher, et al., 2016; Huijser et al., 2020). OPTICS has several advantages over *k*-means clustering, including that points can be designated as outliers rather than forced into clusters, and the number of clusters does not need to be specified *a priori*. However, other input values still need to be set (requiring sensitivity analyses to find initialization values) and significant portions of the data are often designated as outliers (as high as 93% and 70% for certain click lengths in Cantor et al., 2016 and Huijser et al., 2020, respectively), which makes OPTICS extremely conservative and limits the scope and types of questions that can be addressed.

Our ability to delineate clans in an "automated, objective, and meaningful way independent of the dataset" remains a challenging but essential research objective, especially as datasets scale up in size (Frantzis & Alexiadou, 2008). IDcall addresses some of these challenges by focusing on common and distinctive coda types. It takes advantage of two well-supported assumptions: different sperm whale clans make different types of codas (Gero et al., 2016; Rendell & Whitehead, 2003b) and certain coda types are favored by each clan (Amano et al., 2014; Gero et al., 2016; Huijser et al., 2019; Rendell & Whitehead, 2003b). There is theoretical support for this approach (e.g. repetition and redundancy of call types to improve signal-to-noise ratios) if the production of identity codas is more salient to the whales than general repertoire similarity (Wiley, 2013). This is especially the case if some coda types function as 'passwords' or badges of clan membership (e.g. Tyack, 2008), as has been hypothesized by several researchers (Cantor & Whitehead, 2013; Gero, Whitehead, et al., 2016; Whitehead & Rendell, 2014).

One noteworthy deviation of IDcall compared to previously used methods is that call classification is initially probabilistic rather than categorical (i.e. each call is assigned a probability of belonging to each call type rather than being given a single call type designation). The user can eventually choose to categorically assign calls to types (whereby each call is assigned to the type for which is has the highest probability) but does not have to. The user can also choose whether to exclude calls designated as outliers from future analyses.

We compared the coda types detected by IDcall to those detected in previous Atlantic/Mediterranean (Table S3.12) and Pacific (Table S3.13) analyses. Despite variability and inconsistency in coda naming practices across studies, many coda types found in previous Atlantic/Mediterranean and Pacific sperm whale research have similar rhythmic patterns to coda types found using IDcall. We did not attempt to look for 1:1 type matches between IDcall and previous methods given that many past studies determined coda types after standardizing codas for duration (e.g. Rendell & Whitehead, 2003b) or did not provide detailed durational information for each type (e.g. Pavan et al., 2000). Some types were unique to IDcall, which may reflect the previously discussed differences in call classification methodology (many of these unique types are rare and were likely excluded as outliers in past studies) and the increase in sample size in both datasets from previous studies (3,091 new Atlantic/Mediterranean codas; 2,765 new Pacific codas). Additionally, a small number of types were present in past work but not in the present work, which could again relate to sample size increases or variability in naming practices. Importantly, many of the coda types previously recognized as frequent and important in the other studies we consulted were also found by IDcall.

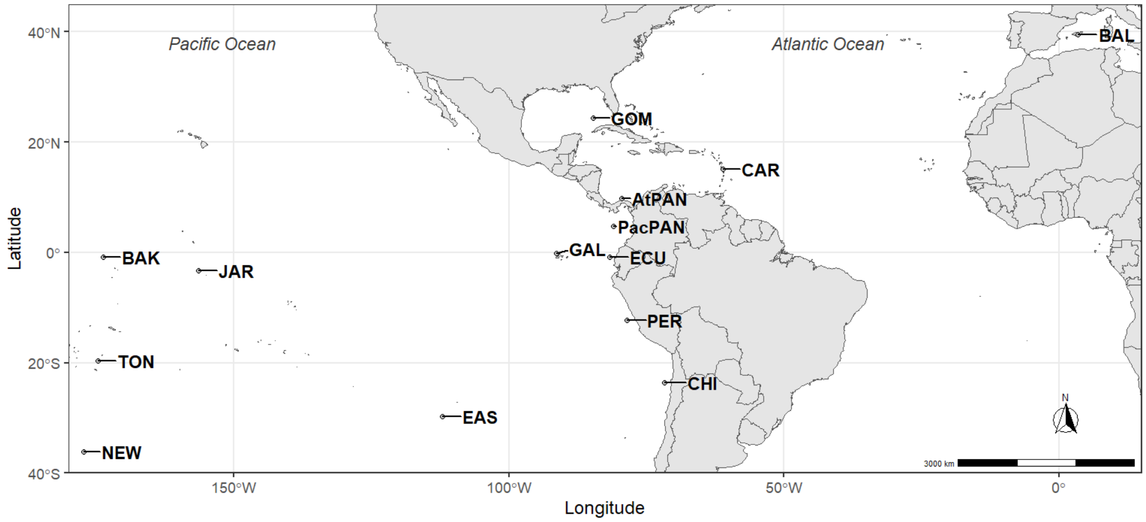
#### *Grey-breasted wood-wrens*

Male *H. l. leucophrys* song is generally slower, contains fewer notes, and is higher and broader in frequency than male *H. l. hilaris* song (Dingle et al., 2008; Halfwerk et al., 2016). Accordingly, the centroid averaged note peak frequency, minimum song frequency, and maximum song frequency were all lower in the *H. l. hilaris* identity song compared to the *H. l. leucophrys* identity songs, matching past work.

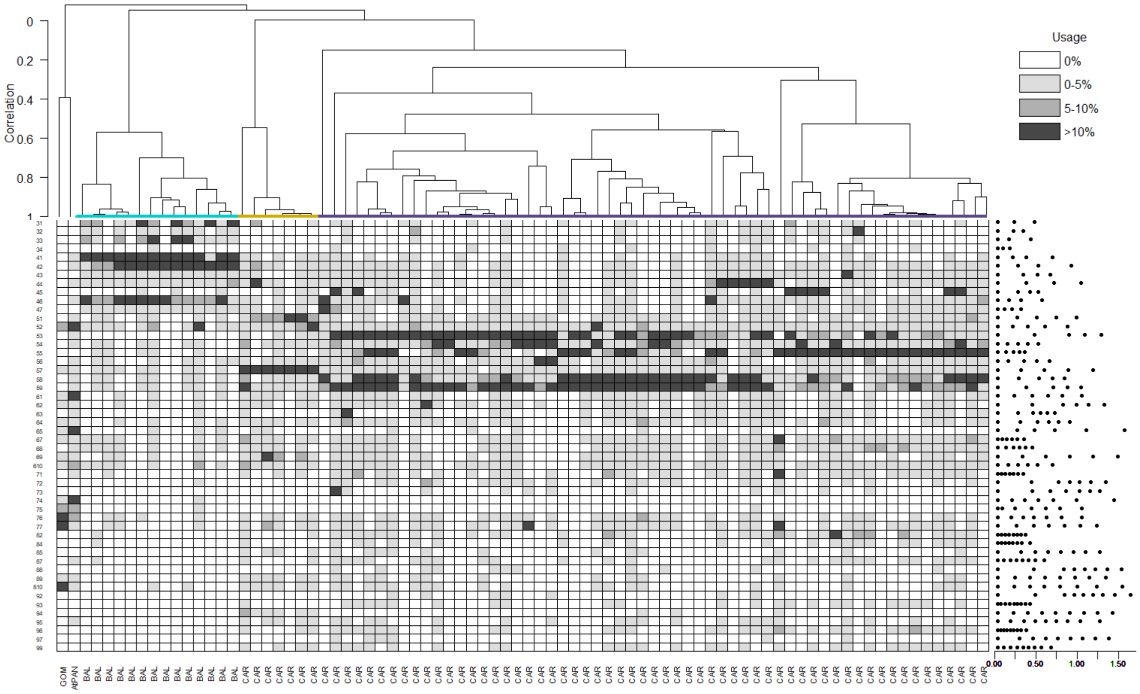
#### *Australian field crickets*

*T. commodus* songs generally have a lower carrier frequency, fewer trills, and a shorter chirp-trill interval than *T. oceanicus* songs (Moran et al., 2020). Accordingly, the centroid chirp-trill interval in the *T. commodus* identity songs was shorter than in the *T. oceanicus* identity songs, matching past work.

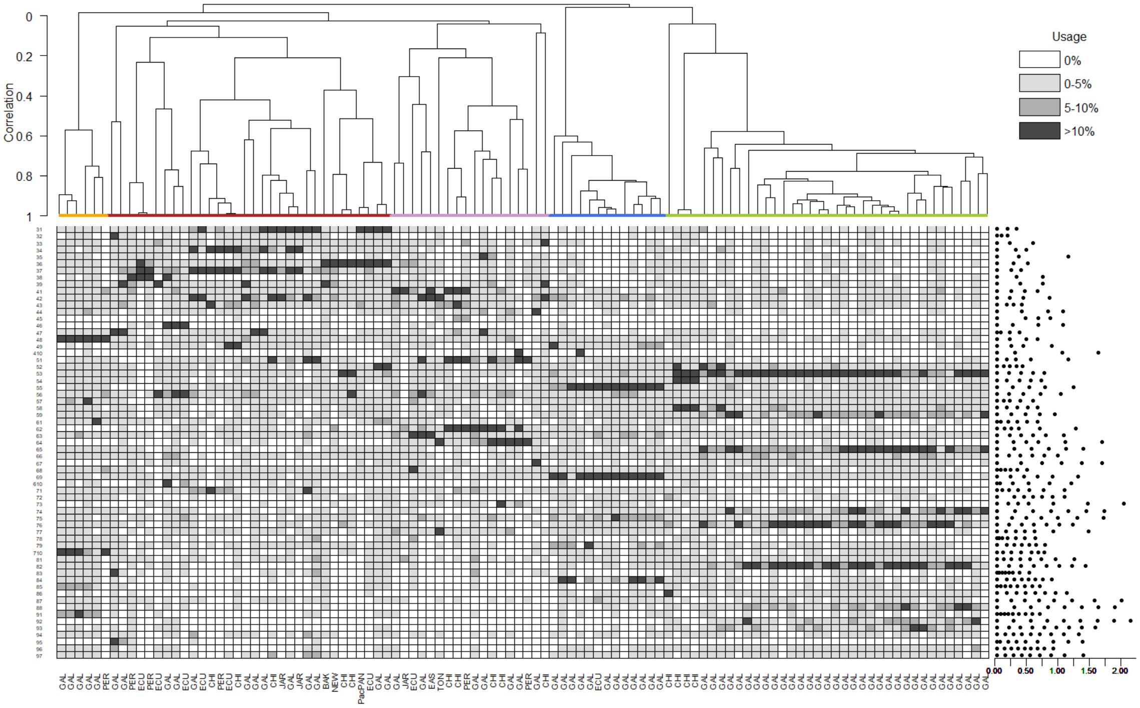
**B3.4:** *Figure S3.1* – Map of sperm whale recording locations. Points on the map show the approximate location of recordings. Pacific abbreviations are: BAK=Baker Island, CHI=Chile, EAS=Easter Island, ECU=Ecuador, GAL=Galápagos Islands, JAR=Jarvis Island, NEW=New Zealand, PacPAN=Pacific coast of Panama, PER=Peru, and TON=Tonga. Atlantic/Mediterranean abbreviations are: AtPAN=Atlantic coast of Panama, BAL=Balearic Islands, CAR=eastern Caribbean, and GOM=Gulf of Mexico.



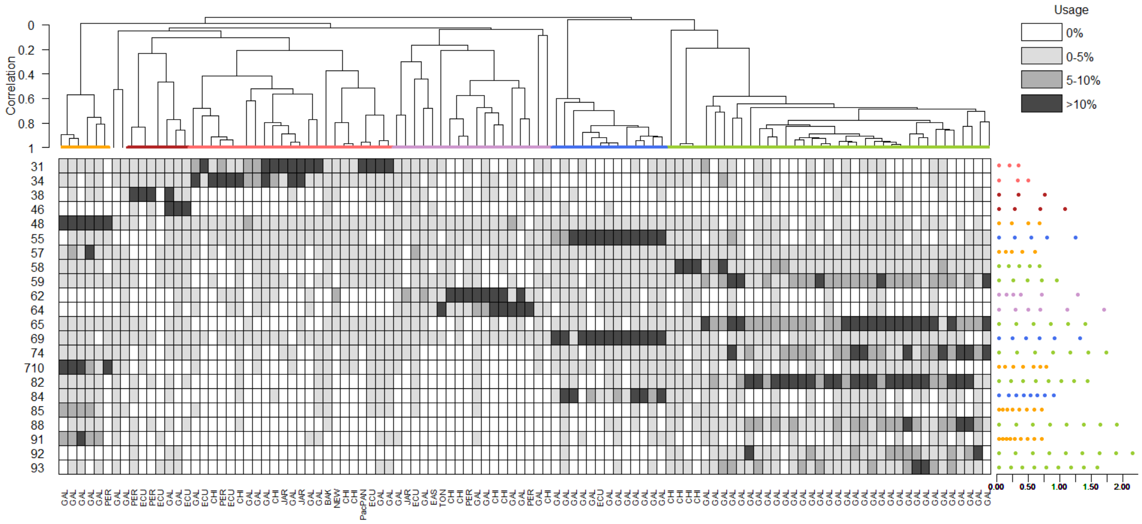
**B3.5:** *Figure S3.2* – Baseline dendrogram with clans and all coda types for Atlantic/Mediterranean sperm whales. Average linkage hierarchical clustering dendrogram (top) depicts similarity among sperm whale coda repertoires recorded in the Atlantic/Mediterranean (see Figure 3.1 for identity clade colors and detailed figure explanation). Heat map (bottom) depicts coda type usage (rows) for each repertoire (columns) (see Figure S3.1 for abbreviations).



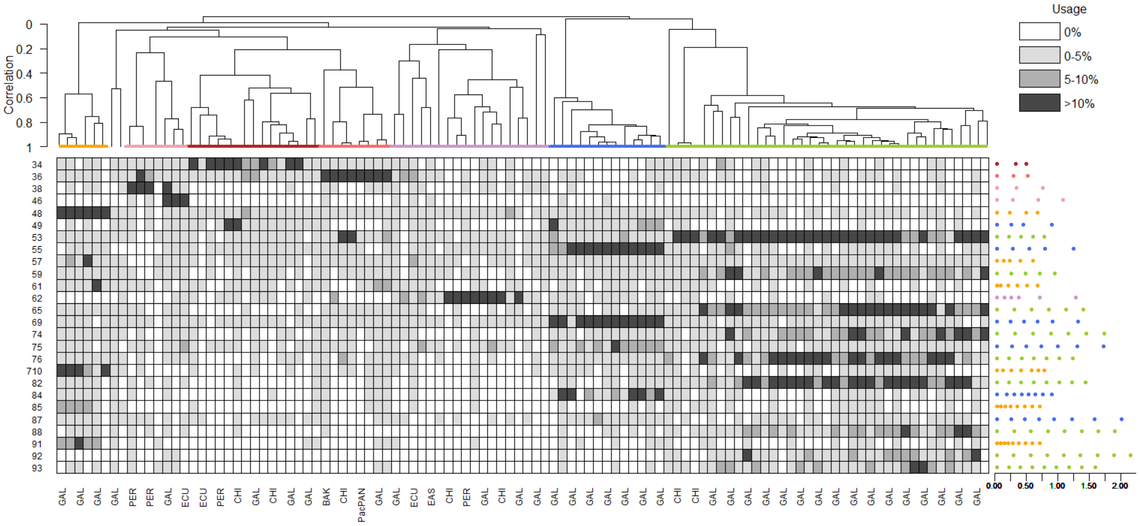
**B3.6:** *Figure S3.3* – Baseline dendrogram with clans and all coda types for Pacific sperm whales. Average linkage hierarchical clustering dendrogram (top) depicts similarity among sperm whale coda repertoires recorded in the Pacific (see Figure 3.2 for identity clade colors and detailed figure explanation). Heat map (bottom) depicts coda type usage (rows) for each repertoire (columns) (see Figure S3.1 for abbreviations).



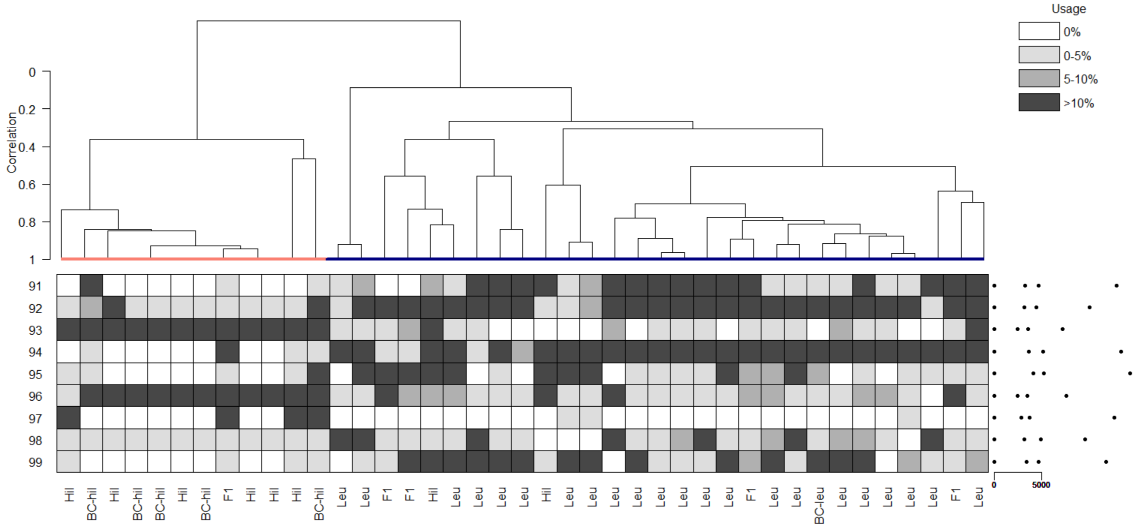
**B3.7:** *Figure S3.4* – Trial dendrogram with clans and identity codas for Pacific sperm whales showing the Short clan divided into two identity clades. Average linkage hierarchical clustering dendrogram (top) depicts similarity among sperm whale coda repertoires recorded in the Pacific. Colored identity clades correspond to a putative new clan (orange), Regular clan (green), Plus-One clan (blue), Four-Plus clan (pink), and two clans (in shades of red) that were considered the Short clan in the baseline dendrogram (Figure 3.2). Heat map (bottom) depicts coda type usage (rows) for each repertoire (columns) (see Figure S3.1 for abbreviations and Figure 3.2 for detailed figure explanation). Dendrogram was created using default parameter values except for *critfact,* which was set to 10. Note that the identity coda type codes here (i.e. the heat map row labels) do not match those in Figures 3.2/S3.3/S3.5 or Table S3.13.



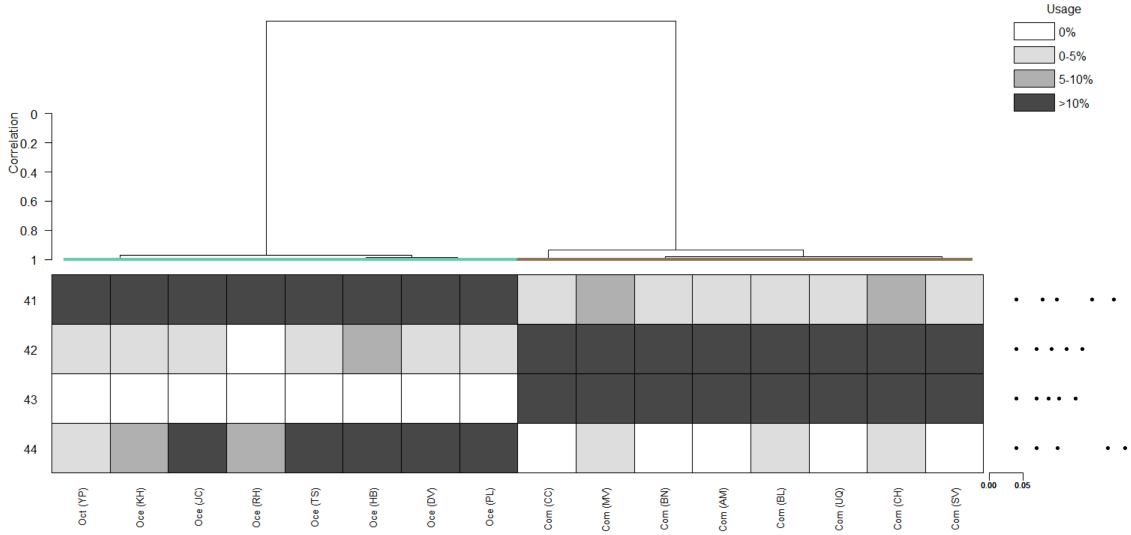
**B3.8:** *Figure S3.5* – Trial dendrogram with clans and identity codas for Pacific sperm whales showing the Short clan divided into three identity clades. Average linkage hierarchical clustering dendrogram (top) depicts similarity among sperm whale coda repertoires recorded in the Pacific. Colored identity clades correspond to a putative new clan (orange), Regular clan (green), Plus-One clan (blue), Four-Plus clan (pink), and three clans (in shades of red) that were considered the Short clan in the baseline dendrogram (Fig. 2). Heat map (bottom) depicts coda type usage (rows) for each repertoire (columns) (see Figure S3.1 for abbreviations and Figure 3.2 for detailed figure explanation). Dendrogram was created using default parameter values except for *critfact,* which was set to 6. Note that the identity coda type codes here (i.e. the heat map row labels) do not match those in Figures 3.2/S3.3/S3.4 or Table S3.13.



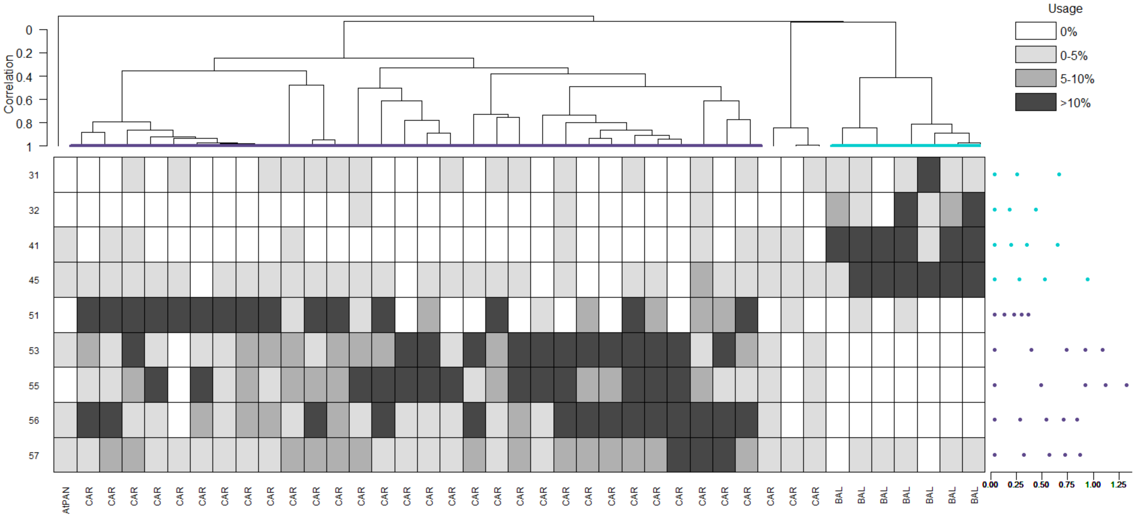
**B3.9:** *Figure S3.6* – Baseline dendrogram with subspecies and all song types for wrens. Average linkage hierarchical clustering dendrogram (top) depicts similarity among repertoires of song frequency vectors of male wrens (see Figure 3.3 for identity clade colors, abbreviations, and detailed figure explanation). Heat map (bottom) depicts song type usage (rows) for each male (columns).



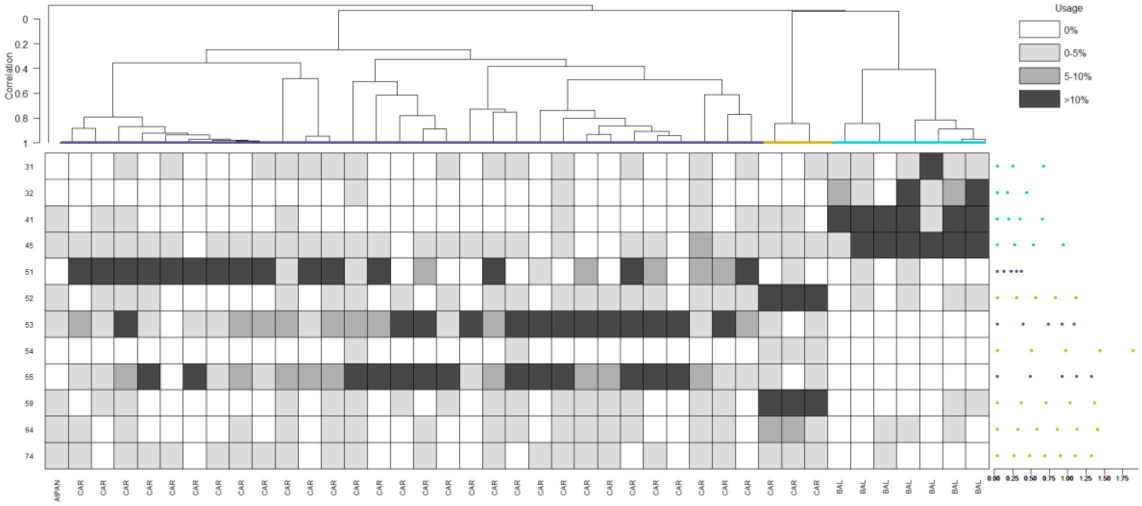
**B3.10:** *Figure S3.7* – Baseline dendrogram with species and all song types for crickets. Average linkage hierarchical clustering dendrogram (top) depicts similarity among repertoires of song interval vectors of male crickets from 16 sites (see Figure 3.4 for identity clade colors, abbreviations, and detailed figure explanation). Heat map (bottom) depicts song type usage (rows) for each field site (columns).



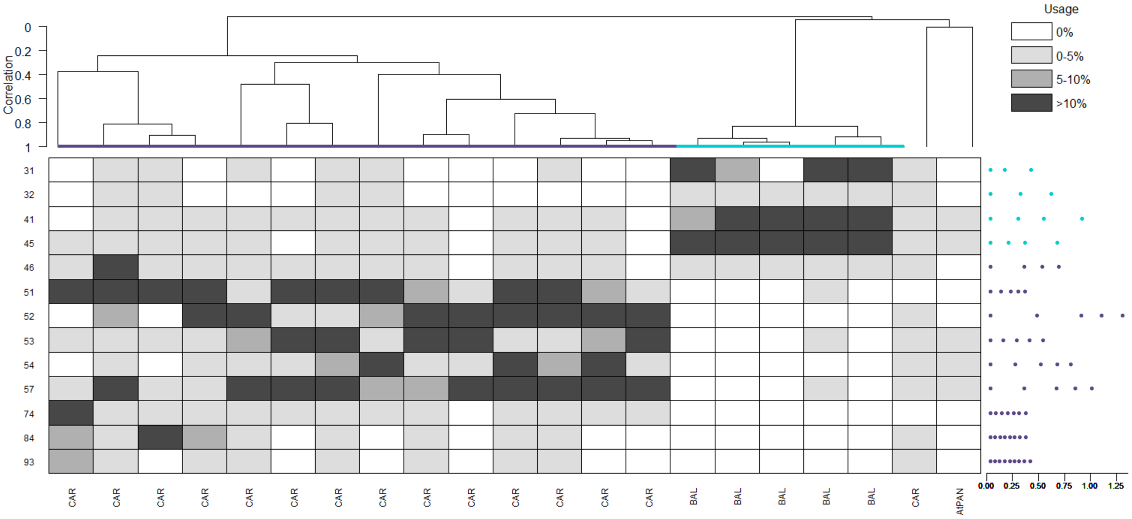
**B3.11:** *Figure S3.8* – Dendrogram with clans and identity coda types for Atlantic/Mediterranean sperm whales using halved dataset and default IDcall values (*critfact*=14, *minrep*=6). Average linkage hierarchical clustering dendrogram (top) depicts similarity among a randomly sampled subset of the sperm whale coda repertoires (n=41, half of the actual dataset) recorded in the Atlantic/Mediterranean (see Figure 3.1 for identity clade colors and detailed figure explanation). The EC2 clan (comprised of three repertoires) does not meet the *minrep* requirement and is not delineated. Heat map (bottom) depicts identity coda type usage (rows) for each repertoire (columns) (see Figure S3.1 for abbreviations). Note that the identity coda type codes here (i.e. the heat map row labels) do not match those in Figures 3.1/S3.2 or Table S3.12.



**B3.12:** *Figure S3.9* – Dendrogram with clans and identity coda types for Atlantic/Mediterranean sperm whales using halved dataset and modified IDcall values (*critfact*=14, *minrep*=3). Average linkage hierarchical clustering dendrogram (top) depicts similarity among a randomly sampled subset of the sperm whale coda repertoires (n=41, half of the actual dataset) recorded in the Atlantic/Mediterranean (see Figure 3.1 for identity clade colors and detailed figure explanation). All three clans are delineated. Heat map (bottom) depicts identity coda type usage (rows) for each repertoire (columns) (see Figure S3.1 for abbreviations). Note that the identity coda type codes here (i.e. the heat map row labels) do not match those in Figures 3.1/S3.2 or Table S3.12.



**B3.13:** *Figure S3.10* – Dendrogram with clans and identity coda types for Atlantic/Mediterranean sperm whales using quartered dataset and modified IDcall values (*critfact*=14, *minrep*=5). Average linkage hierarchical clustering dendrogram (top) depicts similarity among a randomly sampled subset of the sperm whale coda repertoires (n=21, approximately a quarter of the actual dataset) recorded in the Atlantic/Mediterranean (see Figure 3.1 for identity clade colors and detailed figure explanation). The EC2 clan (represented by just one ‘CAR’ repertoire, next to the ‘AtPAN’ repertoire) does not meet the *minrep* requirement and is not delineated. Heat map (bottom) depicts identity coda type usage (rows) for each repertoire (columns) (see Figure S3.1 for abbreviations). Note that the identity coda type codes here (i.e. the heat map row labels) do not match those in Figures 3.1/S3.2 or Table S3.12.



**B3.14:** *Table S3.1* – Atlantic/Mediterranean sperm whale coda datasets. The eastern Caribbean subset contains 10,090 codas recorded between 2005 and 2019 by the Dominica Sperm Whale Project (for field and data collection methods, see Gero et al., 2016; Tønnesen et al., 2018) and 1,551 codas (recorded between 1981 and 1995) extracted from sperm whale master tapes in the Watkins Marine Mammal Sound Database (for field and data collection methods, see Moore et al., 1993; Watkins, 1985; Watkins & Moore, 1982). Eastern Caribbean codas were primarily recorded off Dominica (96.1%), but also off St. Lucia (1.8%), Canouan (1.2%), Bequia (0.6%), and Guadeloupe (0.3%). Coda extraction methods for the Watkins Marine Mammal Sound Database (Sayigh et al., 2016) are summarized in Method S3.1 (Appendix B). Field and data collection methods for the Balearic Islands dataset are detailed in Pirotta et al. (2011) and Rendell et al. (2014). The Gulf of Mexico and Atlantic Panama codas were recorded on single days in 1992 and 1993, respectively (Weilgart & Whitehead, 1997).

| **Recording location** | **Recording years** | **Number of codas (with 3–9 clicks)** | **Number of repertoires** |
|---|---|---|---|
| Eastern Caribbean | 19: 81, 84, 87, 90, 94, 95<br>20: 05, 07–12, 14–16, 18, 19 | 11,641 | 66 |
| Balearic Islands | 20: 04–08, 13, 14, 17, 18 | 1,749 | 14 |
| Panama | 19: 93 | 313 | 1 |
| Gulf of Mexico | 19: 92 | 102 | 1 |
| | **Total:** | **13,805** | **82** |

**B3.15:** *Table S3.2* – Pacific sperm whale coda datasets. A majority of the Pacific codas (18,429) were recorded by our lab between 1985 and 2014 (for field and data collection methods, see Cantor et al., 2016; Rendell & Whitehead, 2003; Weilgart & Whitehead, 1997). Fifty-two codas were recorded in 1978 and extracted from a single sperm whale master tape in the Watkins Marine Mammal Sound Database (Sayigh et al., 2016; Method S3.1 in Appendix B).

| **Recording location** | **Recording years** | **Number of codas (with 3–9 clicks)** | **Number of repertoires** |
|---|---|---|---|
| Galápagos Islands | 19: 78, 85, 87, 89, 91, 95, 99<br>20: 13, 14 | 10,108 | 70 |
| Chile | 19: 93<br>20: 00 | 5,688 | 14 |
| Ecuador | 19: 85, 91, 93 | 773 | 8 |
| Peru | 19: 93 | 657 | 6 |
| Jarvis Island | 19: 92 | 510 | 3 |
| Baker Island | 19: 92 | 269 | 1 |
| Tonga | 19: 92 | 151 | 1 |
| Easter Island | 19: 93 | 90 | 1 |
| Panama | 19: 92 | 191 | 1 |
| New Zealand | 19: 93 | 44 | 1 |
| | **Total:** | **18,481** | **106** |

**B3.16:** *Table S3.3* – Baseline dendrogram metrics for each dataset using the default IDcall parameters.

| **Dataset** | **Number of call types** | **Number of ID call types** | **Number of ID clades** | **Tree ID call proportion** | **Average similarity of trial dendrograms to baseline (± SD)[34]** |
|---|---|---|---|---|---|
| Sperm whales (Atlantic/ Mediterranean) | 56 | 10 | 3 | 0.460 | 0.992 ± 0.017 |
| Sperm whales (Pacific) | 63 | 19 | 5 | 0.378 | 0.972 ± 0.037 |
| Wrens | 9 | 4 | 2 | 0.503 | 0.989 ± 0.032 |
| Crickets | 4 | 4 | 2 | 0.986 | 1.000 ± 0.000 |

---

[34] The average similarity of trial dendrograms to the baseline dendrogram was only calculated for trials that produced a dendrogram (i.e. those trials for which identity calls and identity clades were detected). The number of trials that did not produce a dendrogram for each dataset are: 0 (Atlantic/Mediterranean sperm whales), 0 (Pacific sperm whales), 3 (wrens), and 2 (crickets).

**B3.17:** *Table S3.4* – Atlantic/Mediterranean sperm whale IDcall option/parameter trial results. For each trial, one option/parameter value (the written parameter in each row) was varied while the others were kept at the default/baseline values (top row). Trials 18 and 19 had identical parameters to the baseline.

| **Trial** | **Initialization** | **Criterion** | **Linkage** | ***critfact*** | ***minrep*** | **Number of call types** | **Number of ID call types** | **Tree ID call proportion** | **Number of ID clades** | **Similarity to baseline** |
|---|---|---|---|---|---|---|---|---|---|---|
| Baseline | *k*-means | BIC | average | 14 | 6 | 56 | 10 | 0.460 | 3 | 1.00 |
| 1 | random.post | | | | | 60 | 10 | 0.463 | 3 | 0.996 |
| 2 | random.clas | | | | | 59 | 9 | 0.453 | 3 | 0.996 |
| 3 | | AICc | | | | 60 | 13 | 0.470 | 3 | 1.00 |
| 4 | | AIC | | | | 62 | 13 | 0.460 | 3 | 0.996 |
| 5 | | ICL | | | | 55 | 9 | 0.491 | 3 | 0.982 |
| 6 | | | single | | | 56 | 9 | 0.592 | 3 | 0.929 |
| 7 | | | complete | | | 56 | 7 | 0.677 | 2 | 0.964 |
| 8 | | | | 6 | | 56 | 14 | 0.712 | 3 | 1.00 |
| 9 | | | | 10 | | 56 | 11 | 0.484 | 3 | 1.00 |
| 10 | | | | 18 | | 56 | 9 | 0.453 | 3 | 1.00 |
| 11 | | | | 22 | | 56 | 8 | 0.451 | 3 | 1.00 |
| 12 | | | | 26 | | 56 | 7 | 0.398 | 3 | 1.00 |
| 13 | | | | | 4 | 56 | 10 | 0.460 | 3 | 1.00 |
| 14 | | | | | 8 | 56 | 7 | 0.671 | 2 | 0.996 |
| 15 | | | | | 10 | 56 | 7 | 0.671 | 2 | 0.996 |
| 16 | | | | | 12 | 56 | 7 | 0.671 | 2 | 0.996 |
| 17 | | | | | 14 | 56 | 7 | 0.672 | 2 | 0.996 |
| 18 | | | | | | 56 | 9 | 0.460 | 3 | 0.996 |
| 19 | | | | | | 57 | 11 | 0.471 | 3 | 0.996 |

**B3.18:** *Table S3.5* – Pacific sperm whale IDcall option/parameter trial results. For each trial, one option/parameter value (the written parameter in each row) was varied while the others were kept at the default/baseline values (top row). Trials 18 and 19 had identical parameters to the baseline.

| **Trial** | **Initialization** | **Criterion** | **Linkage** | ***critfact*** | ***minrep*** | **Number of call types** | **Number of ID call types** | **Tree ID call proportion** | **Number of ID clades** | **Similarity to baseline** |
|---|---|---|---|---|---|---|---|---|---|---|
| Baseline | *k*-means | BIC | average | 14 | 6 | 63 | 19 | 0.378 | 5 | 1.00 |
| 1 | random.post | | | | | 65 | 20 | 0.288 | 6 | 0.962 |
| 2 | random.clas | | | | | 63 | 21 | 0.399 | 5 | 0.991 |
| 3 | | AICc | | | | 65 | 19 | 0.374 | 6 | 0.958 |
| 4 | | AIC | | | | 69 | 18 | 0.381 | 5 | 0.988 |
| 5 | | ICL | | | | 61 | 16 | 0.329 | 5 | 0.962 |
| 6 | | | single | | | 63 | 18 | 0.308 | 5 | 0.886 |
| 7 | | | complete | | | 63 | 13 | 0.319 | 5 | 0.872 |
| 8 | | | | 6 | | 63 | 26 | 0.462 | 7 | 0.939 |
| 9 | | | | 10 | | 63 | 22 | 0.392 | 6 | 0.960 |
| 10 | | | | 18 | | 63 | 17 | 0.327 | 5 | 1.00 |
| 11 | | | | 22 | | 63 | 17 | 0.327 | 5 | 1.00 |
| 12 | | | | 26 | | 63 | 14 | 0.264 | 5 | 1.00 |
| 13 | | | | | 4 | 63 | 19 | 0.378 | 5 | 1.00 |
| 14 | | | | | 8 | 63 | 16 | 0.374 | 4 | 1.00 |
| 15 | | | | | 10 | 63 | 16 | 0.374 | 4 | 1.00 |
| 16 | | | | | 12 | 63 | 17 | 0.389 | 4 | 1.00 |
| 17 | | | | | 14 | 63 | 14 | 0.366 | 3 | 0.986 |
| 18 | | | | | | 64 | 18 | 0.361 | 5 | 0.983 |
| 19 | | | | | | 64 | 20 | 0.311 | 6 | 0.957 |

**B3.19:** *Table S3.6* – Wren IDcall option/parameter trial results. For each trial, one option/parameter value (the written parameter in each row) was varied while the others were kept at the default/baseline values (top row). Trials 18 and 19 had identical parameters to the baseline.

| **Trial** | **Initialization** | **Criterion** | **Linkage** | ***critfact*** | ***minrep*** | **Number of call types** | **Number of ID call types** | **Tree ID call proportion** | **Number of ID clades** | **Similarity to baseline** |
|---|---|---|---|---|---|---|---|---|---|---|
| Baseline | *k*-means | AICc | average | 14 | 6 | 9 | 4 | 0.503 | 2 | 1.00 |
| 1 | random.post | | | | | 6 | 2 | 0.307 | 2 | 0.905 |
| 2 | random.clas | | | | | 6 | 0 | 0.000 | 0 | NA |
| 3 | | AIC | | | | 10 | 3 | 0.393 | 2 | 1.00 |
| 4 | | BIC | | | | 3 | 0 | 0.000 | 0 | NA |
| 5 | | ICL | | | | 2 | 0 | 0.000 | 0 | NA |
| 6 | | | single | | | 9 | 4 | 0.503 | 2 | 1.00 |
| 7 | | | complete | | | 9 | 1 | 0.624 | 1 | 0.905 |
| 8 | | | | 6 | | 9 | 4 | 0.519 | 2 | 1.00 |
| 9 | | | | 10 | | 9 | 3 | 0.440 | 2 | 1.00 |
| 10 | | | | 18 | | 9 | 4 | 0.503 | 2 | 1.00 |
| 11 | | | | 22 | | 9 | 3 | 0.305 | 2 | 1.00 |
| 12 | | | | 26 | | 9 | 2 | 0.204 | 1 | 1.00 |
| 13 | | | | | 3 | 9 | 3 | 0.44 | 2 | 1.00 |
| 14 | | | | | 5 | 9 | 3 | 0.44 | 2 | 1.00 |
| 15 | | | | | 7 | 9 | 4 | 0.503 | 2 | 1.00 |
| 16 | | | | | 9 | 9 | 4 | 0.503 | 2 | 1.00 |
| 17 | | | | | 11 | 9 | 4 | 0.503 | 2 | 1.00 |
| 18 | | | | | | 8 | 3 | 0.459 | 2 | 1.00 |
| 19 | | | | | | 9 | 4 | 0.462 | 2 | 1.00 |

**B3.20:** *Table S3.7* – Cricket IDcall option/parameter trial results. For each trial, one option/parameter value (the written parameter in each row) was varied while the others were kept at the default/baseline values (see top row). Trials 18 and 19 had identical parameters to the baseline.

| **Trial** | **Initialization** | **Criterion** | **Linkage** | ***critfact*** | ***minrep*** | **Number of call types** | **Number of ID call types** | **Tree ID call proportion** | **Number of ID clades** | **Similarity to baseline** |
|---|---|---|---|---|---|---|---|---|---|---|
| Baseline | *k*-means | AICc | average | 14 | 6 | 4 | 4 | 0.986 | 2 | 1.00 |
| 1 | random.post | | | | | 3 | 2 | 0.793 | 2 | 1.00 |
| 2 | random.clas | | | | | 3 | 2 | 0.77 | 2 | 1.00 |
| 3 | | AIC | | | | 5 | 3 | 0.891 | 2 | 1.00 |
| 4 | | BIC | | | | 2 | 1 | 0.901 | 1 | 1.00 |
| 5 | | ICL | | | | 2 | 1 | 0.901 | 1 | 1.00 |
| 6 | | | single | | | 4 | 4 | 0.986 | 2 | 1.00 |
| 7 | | | complete | | | 4 | 4 | 0.986 | 2 | 1.00 |
| 8 | | | | 6 | | 4 | 4 | 0.986 | 2 | 1.00 |
| 9 | | | | 10 | | 4 | 4 | 0.986 | 2 | 1.00 |
| 10 | | | | 18 | | 4 | 4 | 0.986 | 2 | 1.00 |
| 11 | | | | 22 | | 4 | 4 | 0.986 | 2 | 1.00 |
| 12 | | | | 26 | | 4 | 4 | 0.986 | 2 | 1.00 |
| 13 | | | | | 3 | 4 | 4 | 0.986 | 2 | 1.00 |
| 14 | | | | | 5 | 4 | 4 | 0.986 | 2 | 1.00 |
| 15 | | | | | 7 | 4 | 4 | 0.986 | 2 | 1.00 |
| 16 | | | | | 9 | 4 | 0 | 0.000 | 0 | NA |
| 17 | | | | | 11 | 4 | 0 | 0.000 | 0 | NA |
| 18 | | | | | | 4 | 4 | 0.985 | 2 | 1.00 |
| 19 | | | | | | 4 | 4 | 0.985 | 2 | 1.00 |

**B3.21:** *Table S3.8* – Identity clades for baseline Atlantic/Mediterranean sperm whale dendrogram.

| **Identity clade name** | **Number of repertoires** | **Average repertoire/clade correlation (± SD)[35]** | **Number of identity calls** | **Clade identity call proportion** |
|---|---|---|---|---|
| EC1 | 59 | 0.652 ± 0.178 | 2 | 0.373 |
| Mediterranean | 14 | 0.849 ± 0.097 | 3 | 0.699 |
| EC2 | 7 | 0.917 ± 0.161 | 5 | 0.717 |

**B3.22:** *Table S3.9* – Identity clades for baseline Pacific sperm whale dendrogram.

| **Identity clade name** | **Number of repertoires** | **Average repertoire/clade correlation (± SD)[35]** | **Number of identity calls** | **Clade identity call proportion** |
|---|---|---|---|---|
| Putative fifth clan | 6 | 0.849 ± 0.088 | 4 | 0.563 |
| Four-Plus | 18 | 0.505 ± 0.286 | 2 | 0.216 |
| Short | 32 | 0.495 ± 0.236 | 3 | 0.321 |
| Plus-One | 13 | 0.872 ± 0.114 | 3 | 0.545 |
| Regular | 37 | 0.781 ± 0.237 | 7 | 0.419 |

**B3.23:** *Table S3.10* – Identity clades for baseline wren dendrogram.

| **Identity clade name** | **Number of repertoires** | **Average repertoire/clade correlation (± SD)[35]** | **Number of identity calls** | **Clade identity call proportion** |
|---|---|---|---|---|
| *H. l. hilaris* | 12 | 0.844 ± 0.228 | 1 | 0.548 |
| *H. l. leucophrys* | 29 | 0.632 ± 0.288 | 3 | 0.485 |

**B3.24:** *Table S3.11* – Identity clades for baseline cricket dendrogram.

| **Identity clade name** | **Number of repertoires** | **Average repertoire/clade correlation (± SD)[35]** | **Number of identity calls** | **Clade identity call proportion** |
|---|---|---|---|---|
| *T. oceanicus* | 8 | 0.991 ± 0.013 | 2 | 0.993 |
| *T. commodus* | 8 | 0.987 ± 0.021 | 2 | 0.978 |

---

[35] Calculated as the average correlation of each repertoire's type usage with the clade's type usage (i.e. the median type usage of all repertoires in the clade).

**B3.25:** *Table S3.12* – Atlantic/Mediterranean sperm whale coda types from IDcall and past work. Method S3.3 (Appendix B) contains some key details for interpreting this Table. Type names refer to the overarching rhythmic pattern of clicks in each coda regardless of duration, which is why multiple numeric codes can be linked to one type. ‘R’ stands for ‘Regular’ (i.e. all inter-click intervals (ICIs) are approximately equal), ‘+’ denotes an extended pause between clicks, ‘D’ stands for ‘Decreasing’ (i.e. ICIs become shorter throughout the coda), and ‘I’ stands for ‘Increasing’ (i.e. ICIs become longer throughout the coda). To name the coda types detected by IDcall, the centroid ICI values for each type were divided by the smallest ICI and the resultant values were inspected to determine the overall rhythmic pattern of the coda. Values within ~0.3 of each other were considered the same for naming purposes. Bolded numeric codes are identity codas and are colored by clan (see Figure 3.1 for colors). The final four columns show coda types detected in past work using different methods that *rhythmically* match coda types found using IDcall (see Method S3.3 in Appendix B and the publications themselves for additional details on how the coda types were delineated and named). The final row shows the number of coda types from each study that rhythmically match coda types in the present study over the total number of 3- to 9-click coda types in each study.

| **Number of clicks** | **Numeric code(s)** | **Type name** | **Gero, Bøttcher, et al. (2016)** | **Gero, Whitehead, et al. (2016)** | **Pavan et al. (2000)** | **Drouot et al. (2004)[36]** |
|---|---|---|---|---|---|---|
| 3 | 31, 34 | 3R | 3R | 3R | | 3reg |
| | **33** | 2+1 | | | | |
| | 32 | 1+2 | 3D | 3D | | |
| 4 | **41**, **42**, 46 | 3+1 | | | 3+1 | 3+1 |
| | 43 | 1+3 | $1+3_1$, $1+3_2$ | $1+3_1$, $1+3_2$ | | |
| | 44, 47 | 4R | 4R1, 4R2 | 4R1, 4R2 | | 4reg |
| | 45 | 4D | 4D | 4D | | |
| 5 | **53**, 58, 59 | 1+1+3 | 1+1+3 | 1+1+3 | | |
| | **51**, 54, 56, **57** | 5R | 5R1, 5R2, 5R3 | 5R1, 5R2, 5R3 | | 5reg |
| | 52 | 4+1 | | | 4+1 | |
| | **55** | 5D | | | | |
| 6 | 61, 68, **69** | 6R | 6R | | | 6reg |
| | 62, 63, 64 | 1+1+4 | | 6D | | |
| | 65 | 4+1+1 | | | | |
| | 66 | 2+1+1+1+1 | | | | |
| | 67, 610 | 5+1 | 6I | 6I | | |

*Continued on next page*

---

[36] We compared coda types detected using IDcall to those recorded by Drouot et al. (2004) in the southwestern basin of the Mediterranean Sea since the codas included in the present study were recorded off the Balearic Islands. The Drouot et al. (2004) study also includes codas recorded in the Tyrrhenian and Ionian Seas.

| Number of clicks | Numeric code(s) | Type name | Gero, Bøttcher, et al. (2016) | Gero, Whitehead, et al. (2016) | Pavan et al. (2000) | Drouot et al. (2004) |
|---|---|---|---|---|---|---|
| 7 | 71, 77 | 7R | 7R | | | 7reg |
| | 72 | 1+1+5 | | | | |
| | 73 | 7D | 7D | 7D | | |
| | 74 | 5+1++1 | | | | |
| | 75 | 2+1+1+1+1+1 | | | | |
| | 76 | 6+1 | 7I | 7I | | |
| 8 | 81 | 6+1+1 | | | | |
| | 82, 87, **89** | 8R | | | | |
| | 83 | 4+4 | | | | |
| | 84, 810 | 7+1 | 8I | 8I | | |
| | 85 | 1+7 | | 8D | | |
| | 86, 88 | 1+1+6 | | | | |
| 9 | 91 | 6+1+1+1 | | | | |
| | 92 | 1++1++7 | | | | |
| | 93, **94** | 9R | 9R | | | |
| | 95, 96, 99 | 8+1 | 9I | 9I | | |
| | 97 | 1+8 | | 9D | | |
| | 98 | 9I | | | | |
| **Type matches/total types found:** | | | **19/20**[37] | **19/20**[38] | **2/2** | **6/7**[39] |

---

[37] Missing type: 2+3

[38] Missing type: 2+3

[39] Missing type: 3++1. There were additional coda types with "undefined patterns" (3var, 4var, 5var, 6var) detected in the southwestern basin of the Mediterranean Sea, but we were unable to determine if these types were also detected by IDcall because the type patterns/durations were not described in Drouot et al. (2004). Those types are not included in the total coda type count.

**B3.26:** *Table S3.13* – Pacific sperm whale coda types from IDcall and past work. See Table S3.12 for details and Figure 3.2 for colors.

| Number of clicks | Numeric code(s) | Type name | Weilgart & Whitehead (1997) | Rendell & Whitehead (2003) | Cantor et al. (2016) |
|---|---|---|---|---|---|
| 3 | 32, 33, 35, 38 | 2+1 | 3a, 2+1 | 2+1 | 2+1 |
| | **34**, **37** | 1+2 | 1+2 | 1+2 | 1+2 |
| | **31**, 36, 39 | 3R | 3R, 3b | 3R | 3R |
| 4 | 46 | 2+1+1 | | | |
| | 47 | 4I | 4L | | |
| | 42, 49 | 3+1 | 3+1 | 3+1 | 3+1 |
| | 41, 43 | 4R | 4R | 4R | 4R |
| | 44, **48** | 1+1+2 | | 2+2 | |
| | 45 | 1+3 | 4Var example | 1+3 | |
| | 410 | 2++1+1 | | | |
| 5 | **55** | 4+1 | 4+1 | 4+1 | 4+1 |
| | 52, 53, 54, **58**, 59 | 5R | 5R | 5R | 5R |
| | 56 | 5I | | | |
| | 51, 57 | 3+1+1 | | | |
| 6 | 61 | 2+1+1+1+1 | 2+4 | | 2+4 |
| | **62**, **64** | 4+1++1 | 4+1++++1 | 4+1+1 | 4+1+1 |
| | 63, **69** | 5+1 | 5+1, 6Var example | 5+1 | 5+1 |
| | **65**, 66, 67 | 6R | 6R | 6R | 6R1, 6R2 |
| | 68, 610 | 6I | | | 6I |
| 7 | **74**, 76, 79 | 7R | 7R, 7Var example | 7R | 7R1, 7R2 |
| | 71, 78 | 7I | | | |
| | 72 | 1+6 | | 1+6 | |
| | 73 | 4+1++1++1 | | 4+3 | |
| | 75 | 6+1 | 6+1 | 6+1 | |
| | 77 | 5+1++1 | 5+1++1 | 5+2A, 5+2B | |
| | **710** | 3+1+3 | | | |
| 8 | **82**, **84**, 86, **88** | 8R | 8R, 8L | 8RA, 8RB | 8R1, 8R2 |
| | 81, 83, **85** | 8I | 8Var example | 8S | |
| | 87 | 6+1+1 | | 6+1+1 | |
| 9 | **91**, 94, 95, 97 | 9I | | | |
| | **92**, **93**, 96 | 9R | | | 9R1, 9R2, 9R3 |
| **Type matches/total types found:** | | | **18/22**[40] | **25/33**[41] | **18/20**[42] |

---

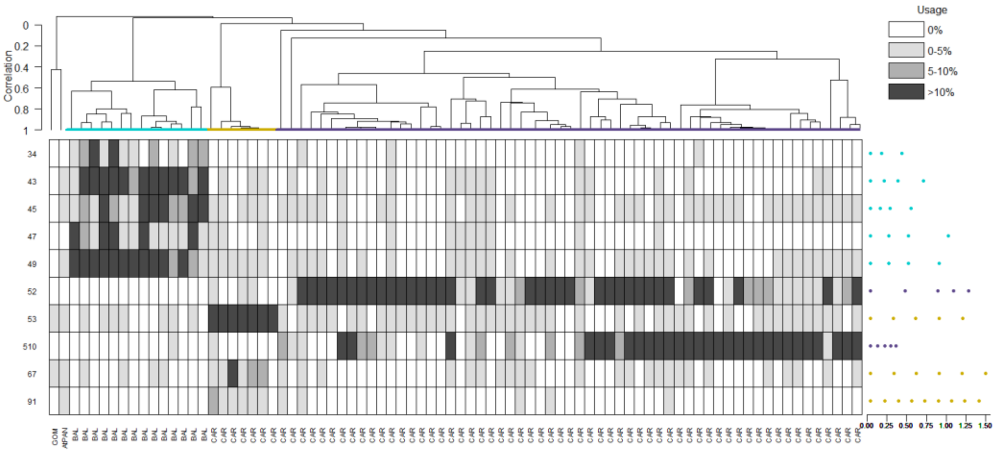
[40] 'Var' coda types are not included in the total coda type count because they contain all coda type clusters that contained less than 50 codas, and likely represent more than one type. Missing types: 3++1, 4++1, 2+1+1+1, and 9.

[41] Rendell & Whitehead (2003) did not analyze 9-click codas. Missing types: 4A, 3++1, 4++1, 3+4, 7+1, 8A, and 7++1.

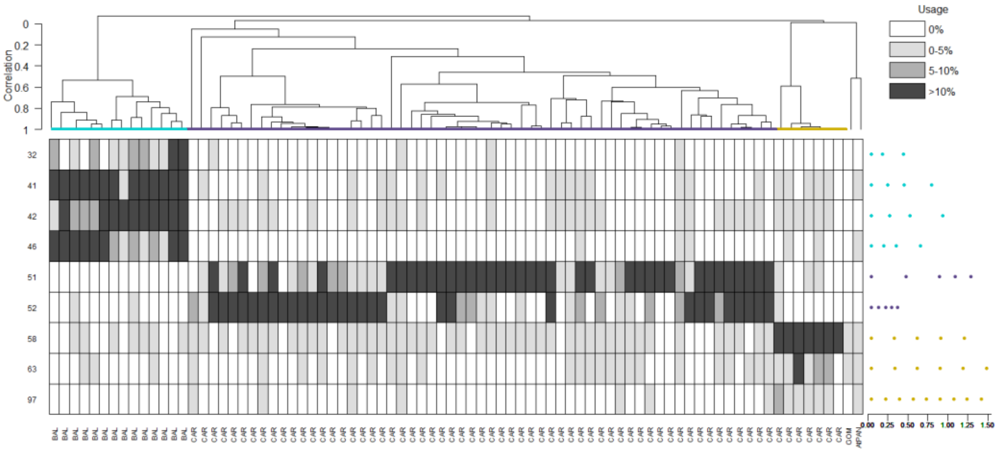
[42] Missing types: 1+2+1, 1+3+1

**B3.27:** *Data S3.1* – Atlantic/Mediterranean sperm whale trial results. The resultant dendrogram for each trial (n=19) is presented below, with trial numbers corresponding to those in Table S3.4. The varied option/parameter is listed after each trial number. All other options/parameters were kept at the default values (Table 3.2). For detailed figure description, see Figure 3.1. Acoustic clade colors match those used in Figure 3.1. Numeric call type codes (i.e. heat map row labels) are consistent for the baseline dendrogram and the trials varying the hierarchical clustering linkage strategy, *critfact*, and *minrep*, but not necessarily for the trials varying the ECM algorithm initialization strategy, the information criterion used during call classification, or the default value reruns.

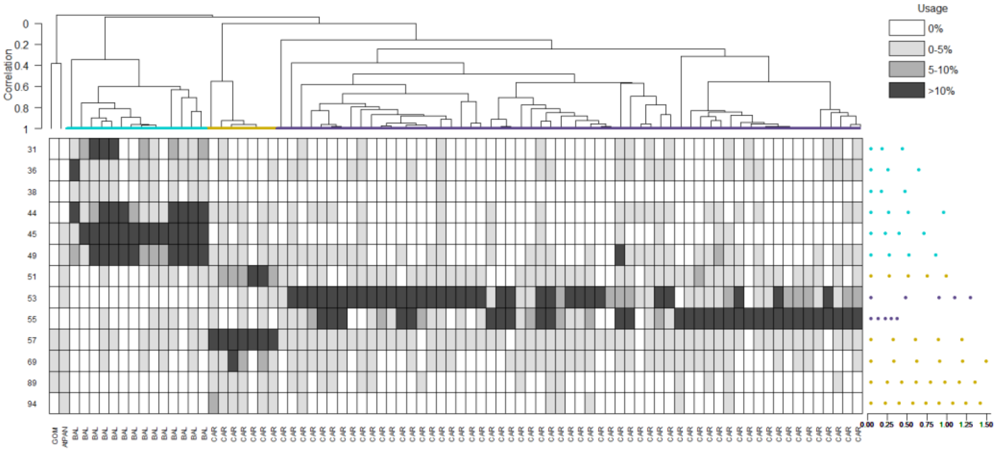
*Trial 1 (initialization=random.post)*



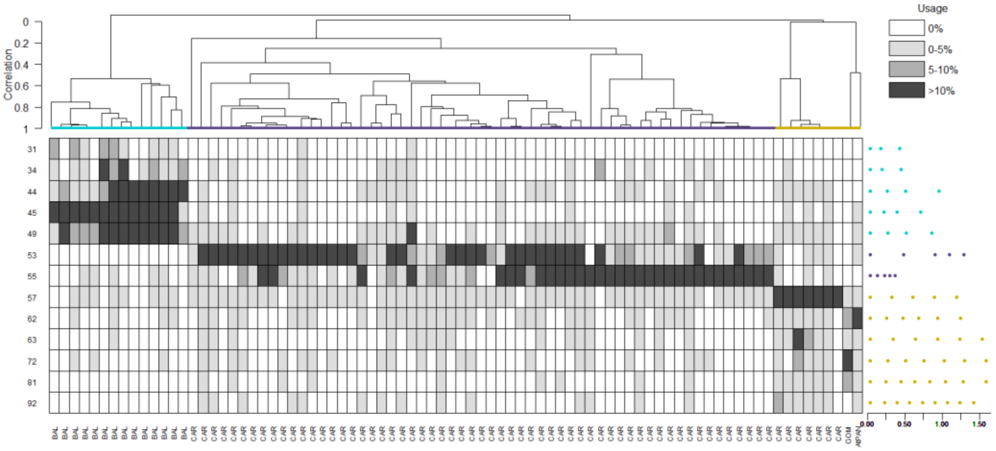
*Trial 2 (initialization=random.clas)*



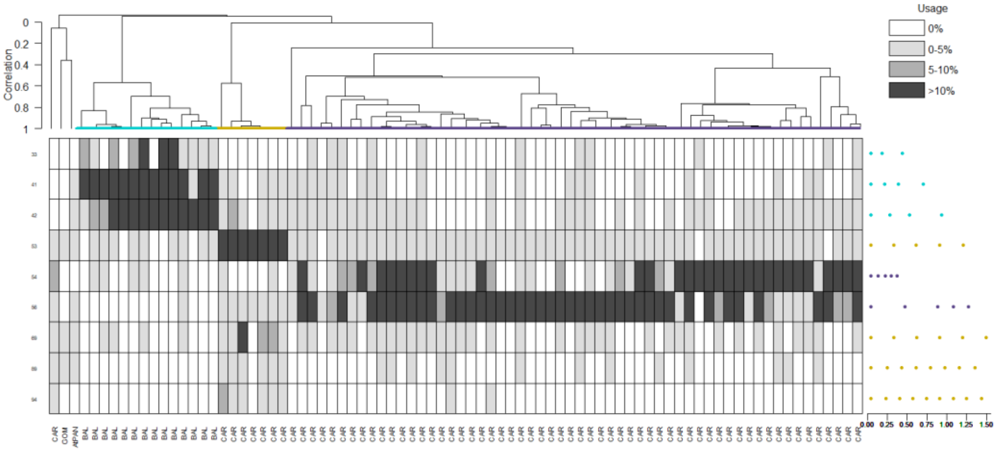
## *Trial 3 (information criterion=AICc)*



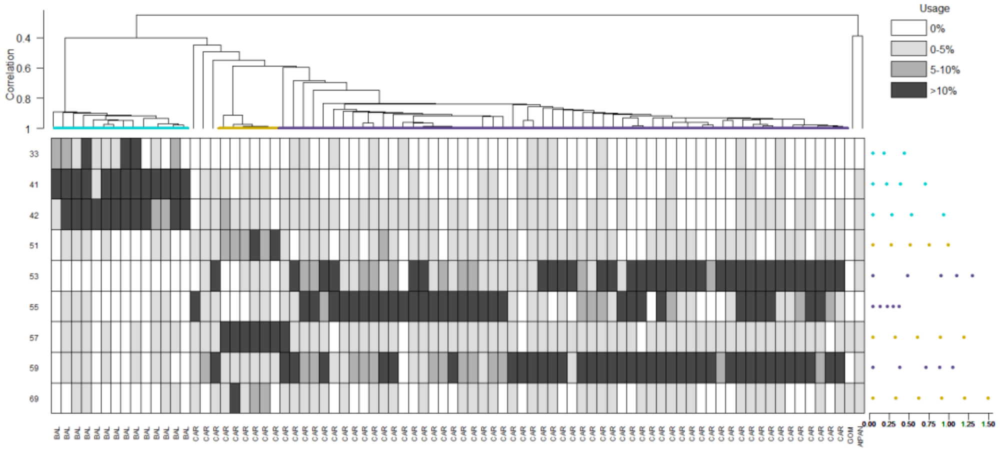
## *Trial 4 (information criterion=AIC)*



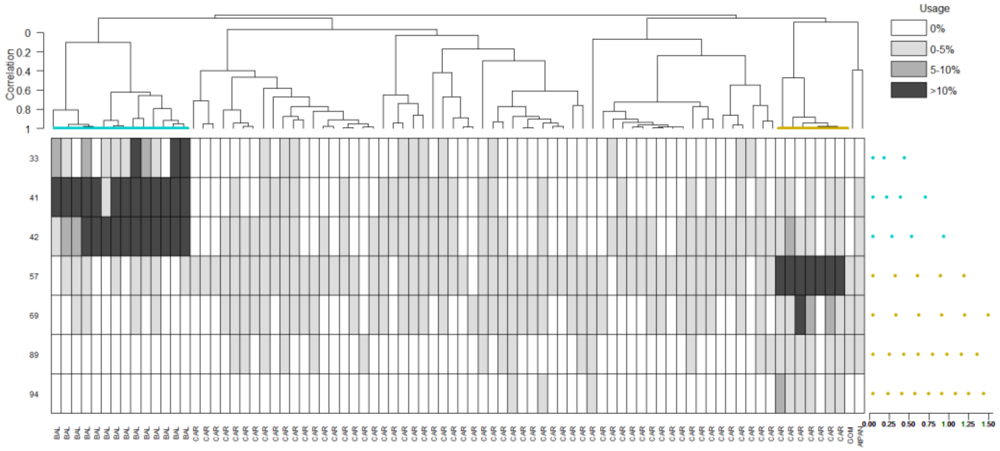
## *Trial 5 (information criterion=ICL)*



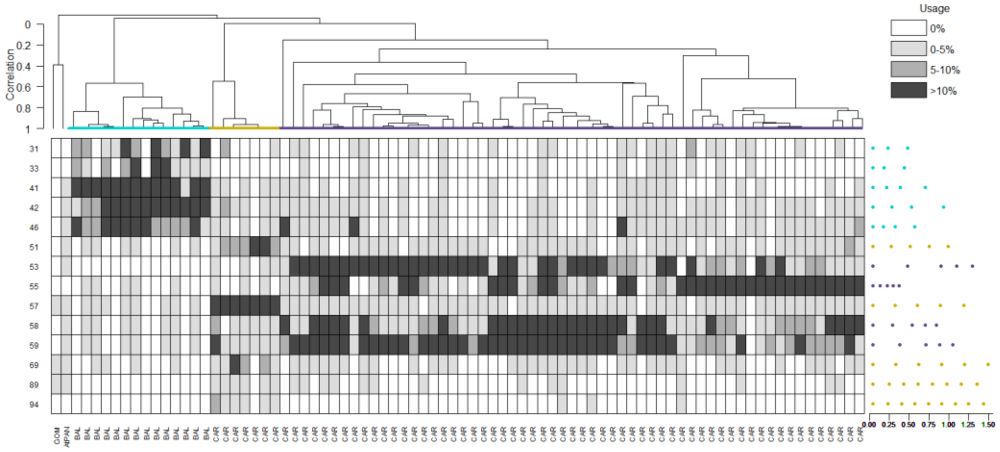
### *Trial 6 (linkage=single)*



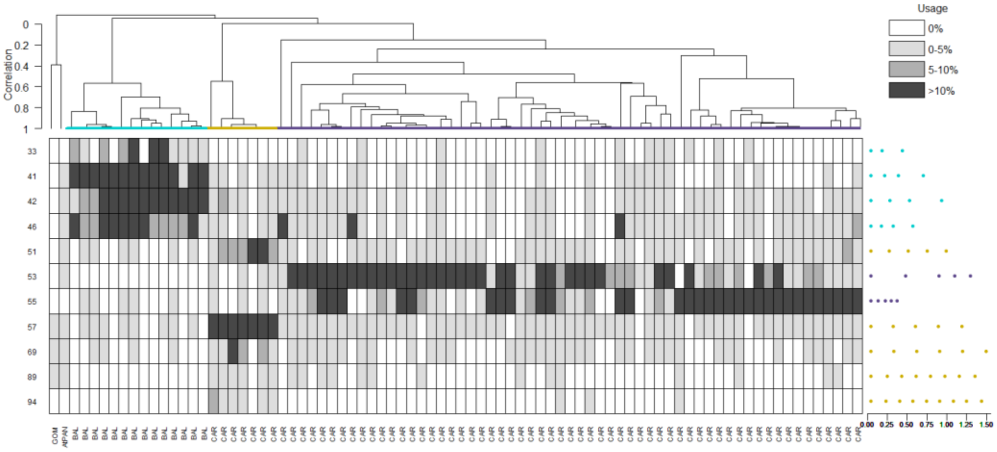
### *Trial 7 (linkage=complete)*



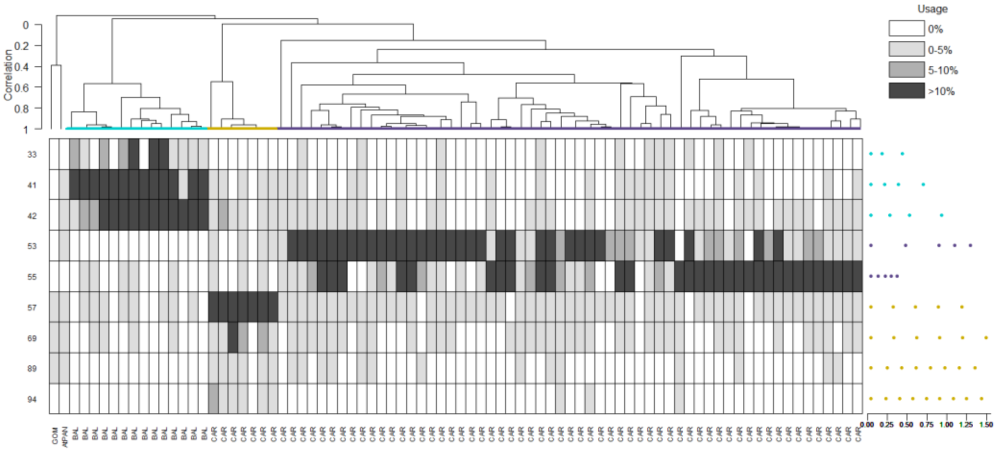
### *Trial 8 (critfact=6)*



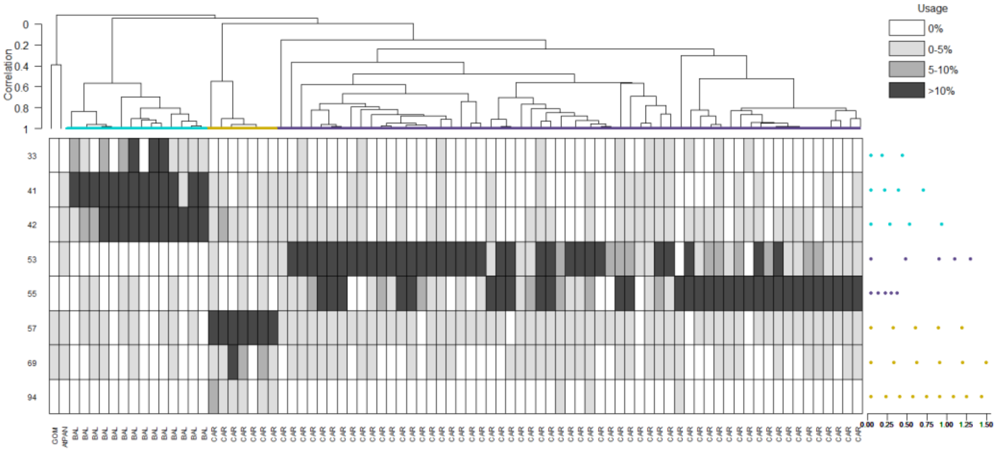
*Trial 9 (critfact=10)*



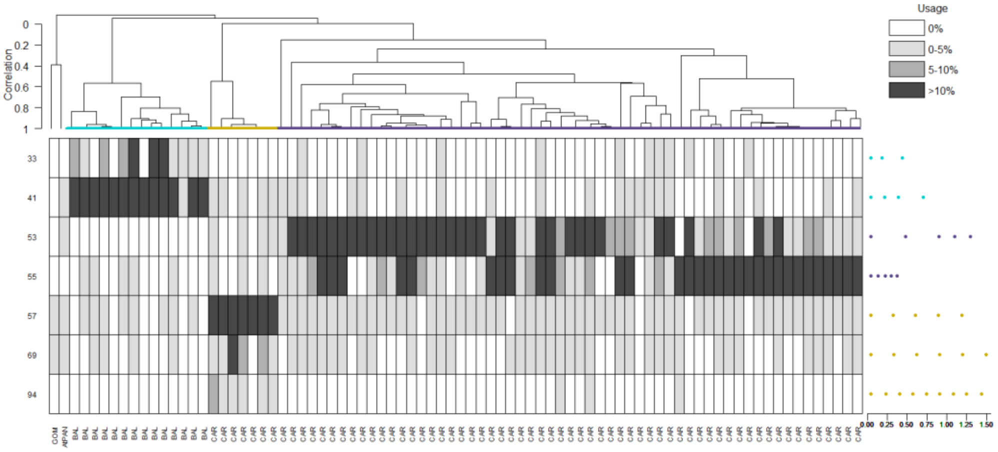
*Trial 10 (critfact=18)*



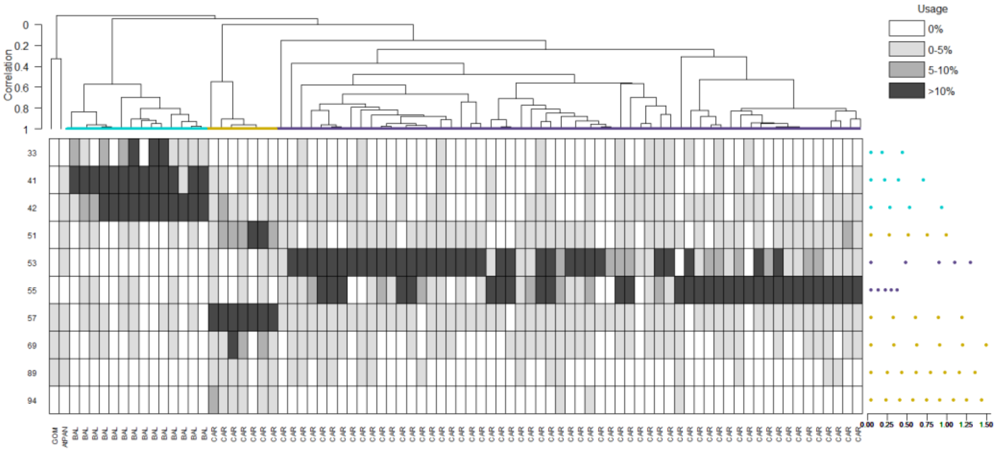
*Trial 11 (critfact=22)*



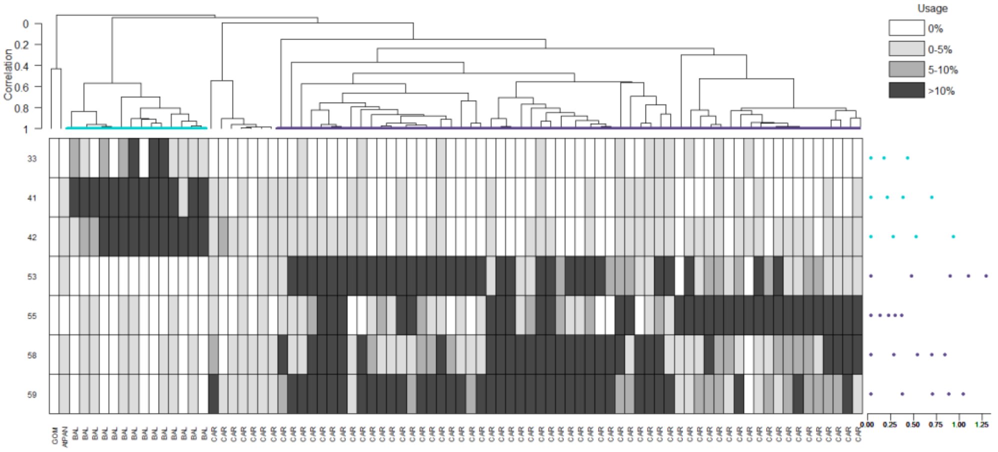
*Trial 12 (critfact=26)*



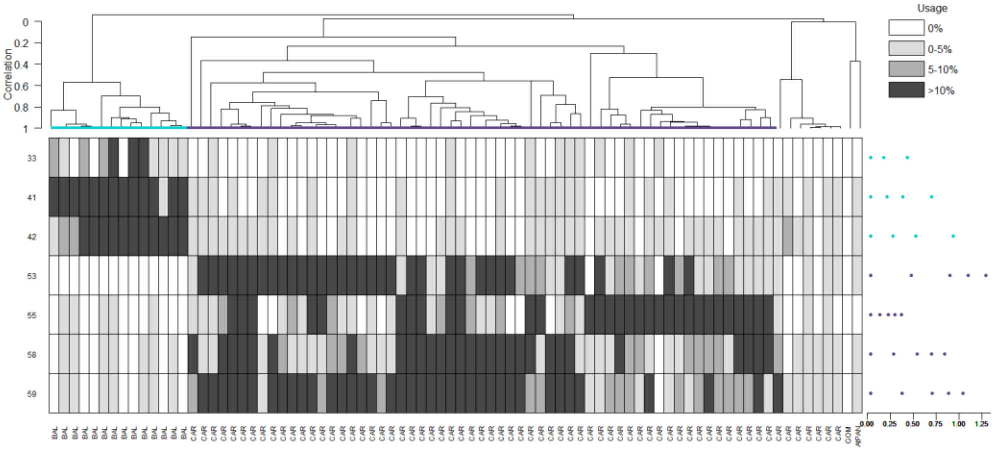
*Trial 13 (minrep=4)*



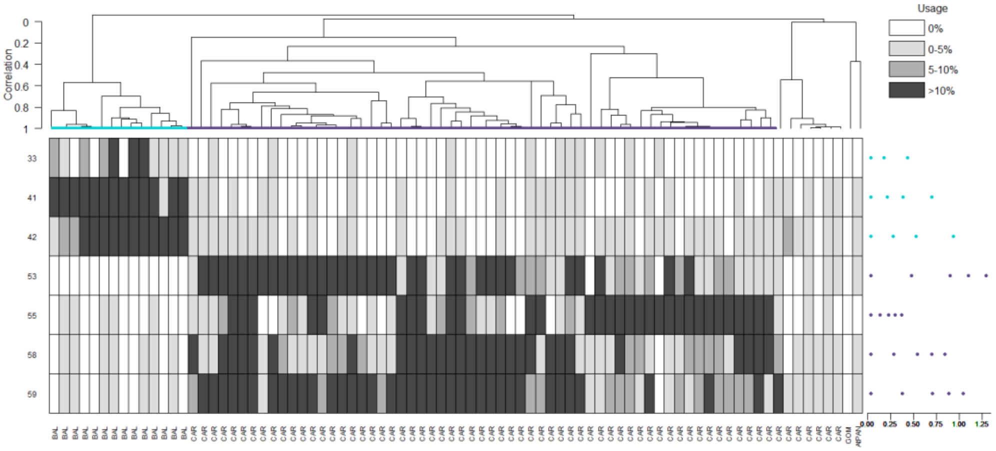
*Trial 14 (minrep=8)*



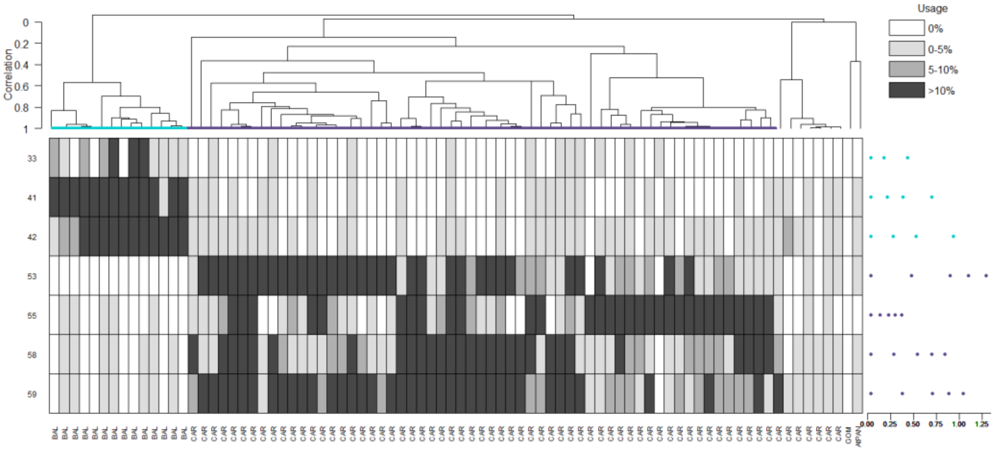
*Trial 15 (minrep=10)*



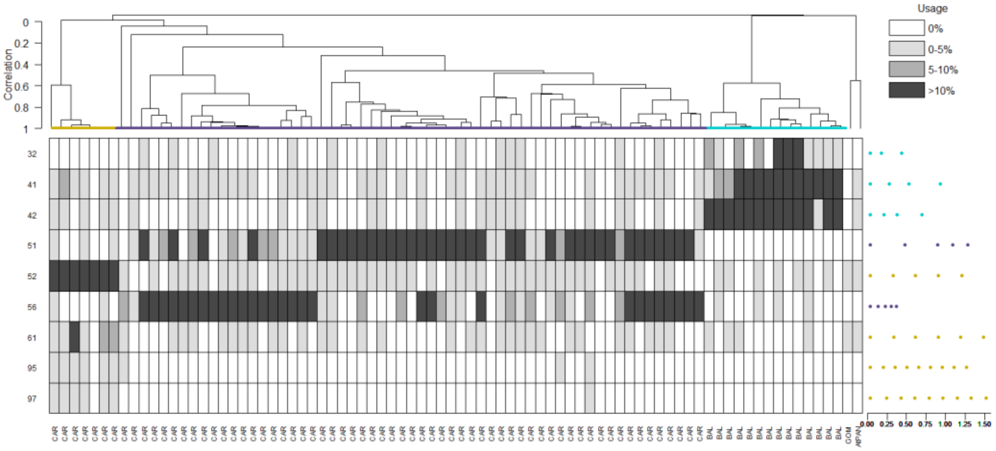
*Trial 16 (minrep=12)*



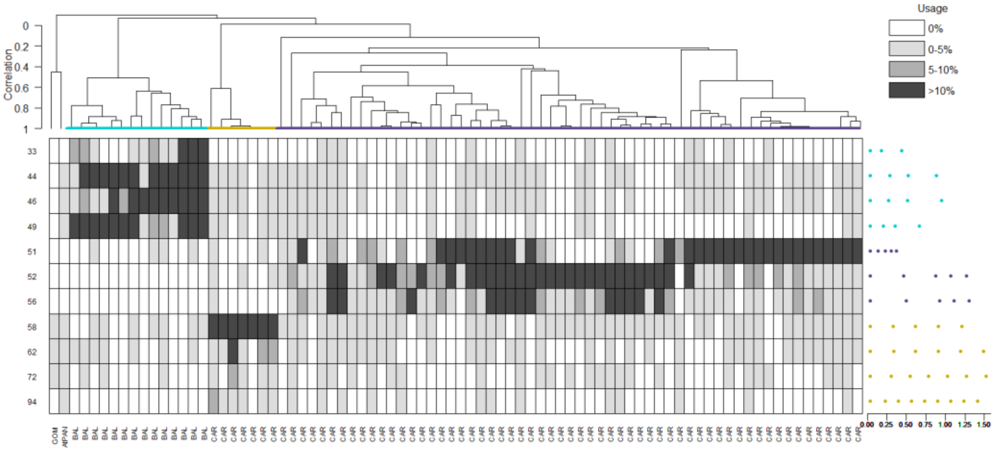
*Trial 17 (minrep=14)*



*Trial 18 (baseline run 2)*

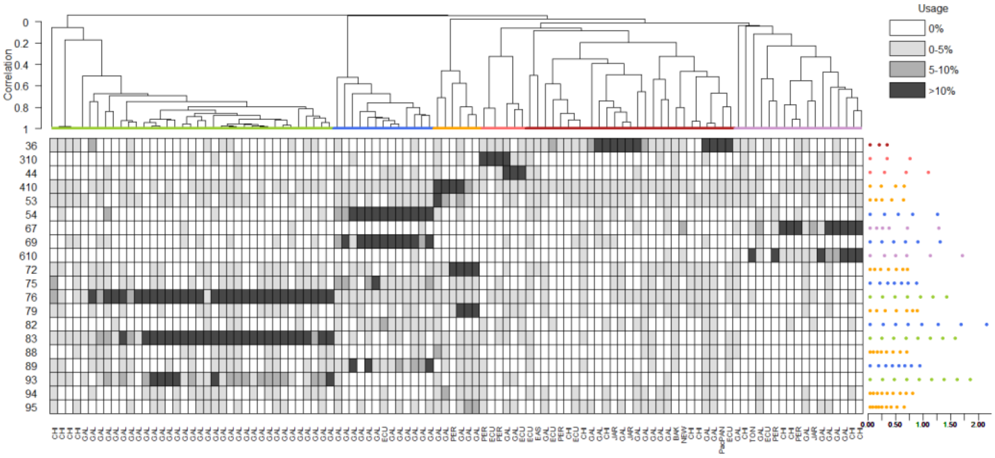


*Trial 19 (baseline run 3)*

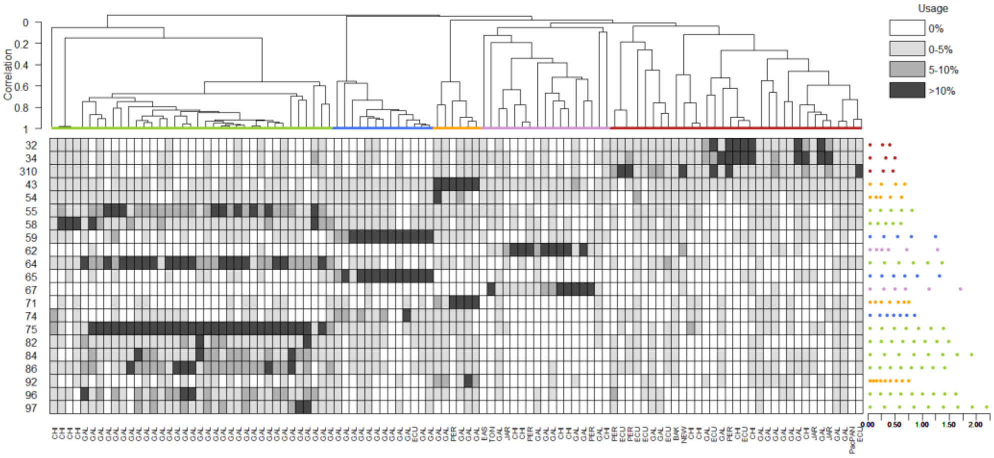


**B3.28:** *Data S3.2* – Pacific sperm whale trial results. The resultant dendrogram for each trial (n=19) is presented below, with trial numbers corresponding to those in Table S3.5. The varied option/parameter is listed after each trial number. All other options/parameters were kept at the default values (Table 3.2). For detailed figure description, see Figure 3.2. Acoustic clade colors match those used in Figures 3.2/S3.4/S3.5. Numeric call type codes (i.e. heat map row labels) are consistent for the baseline dendrogram and the trials varying the hierarchical clustering linkage strategy, *critfact*, and *minrep*, but not necessarily for the trials varying the ECM algorithm initialization strategy, the information criterion used during call classification, or the default value reruns.

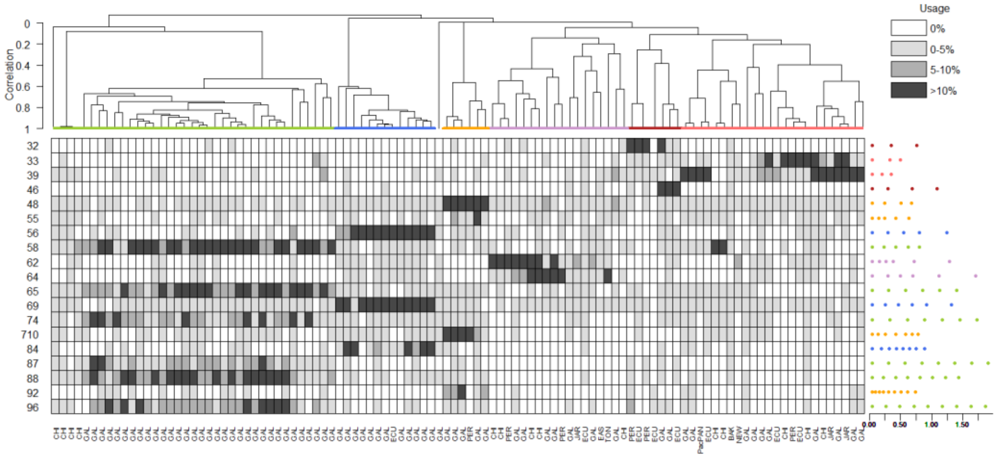
*Trial 1 (initialization=random.post)*



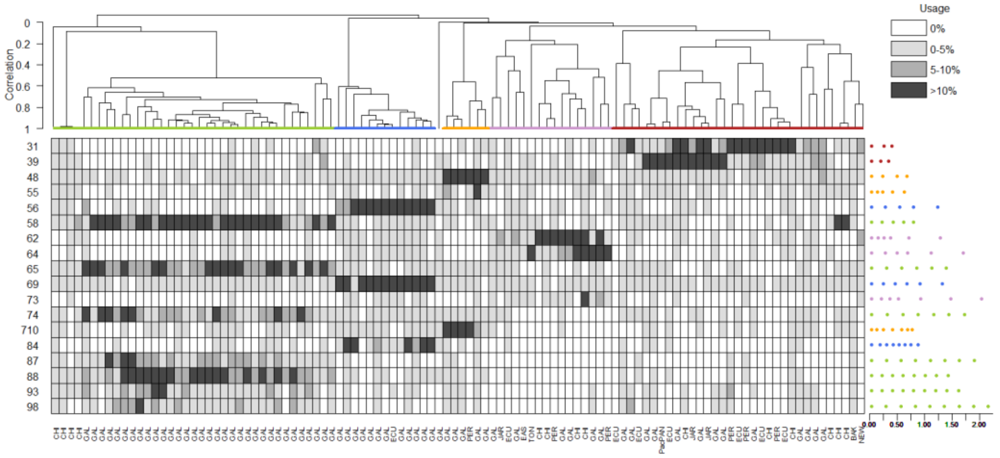
*Trial 2 (initialization=random.clas)*



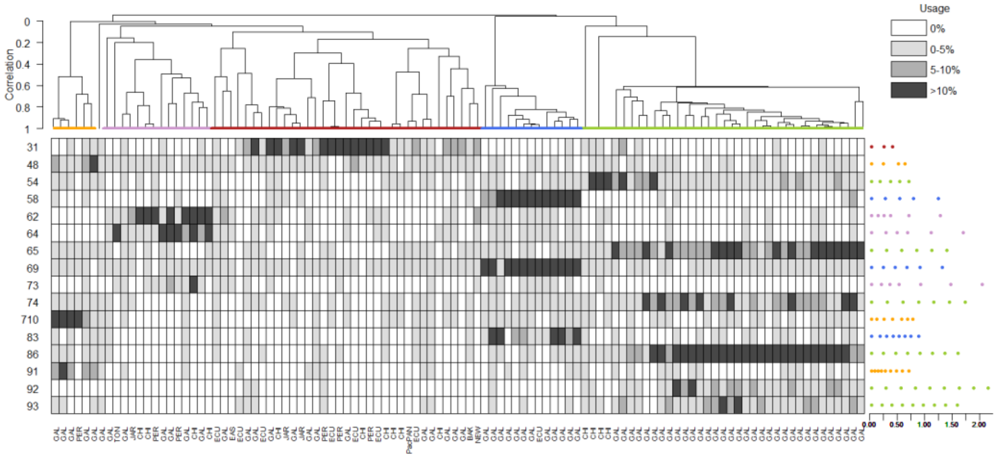
### *Trial 3 (information criterion=AICc)*



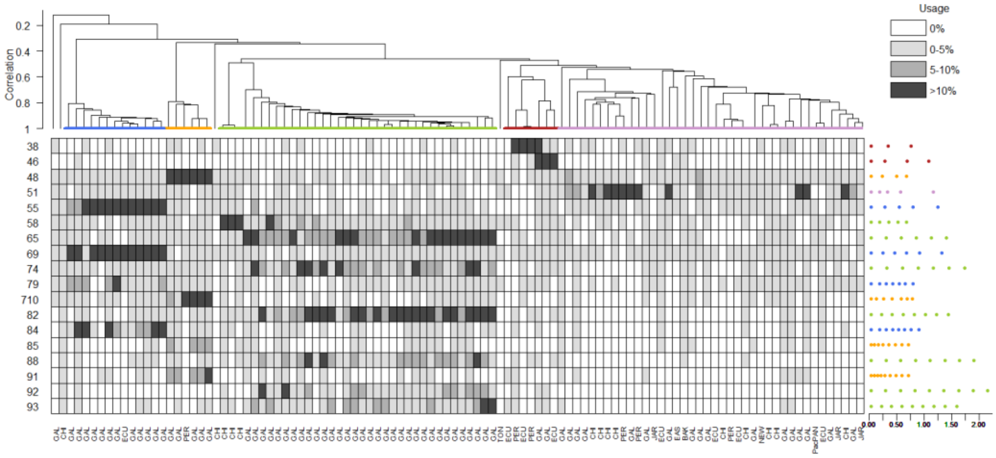
### *Trial 4 (information criterion=AIC)*



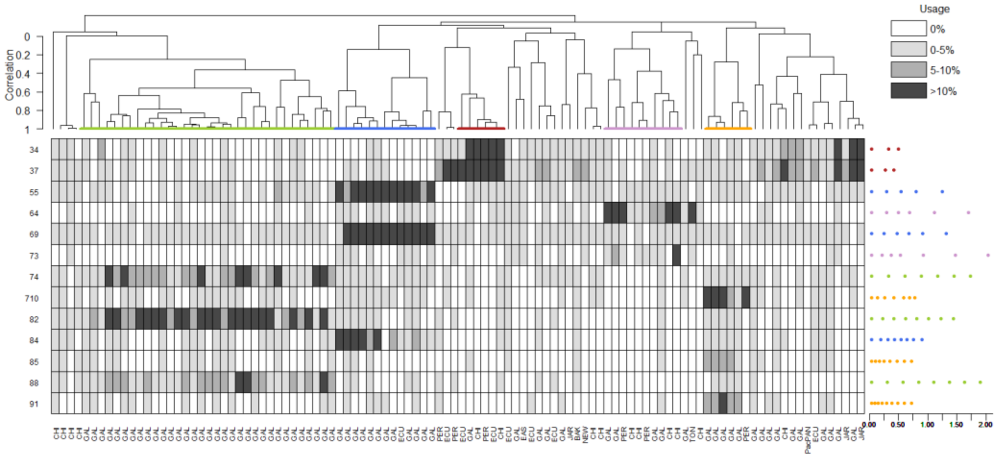
### *Trial 5 (information criterion=ICL)*



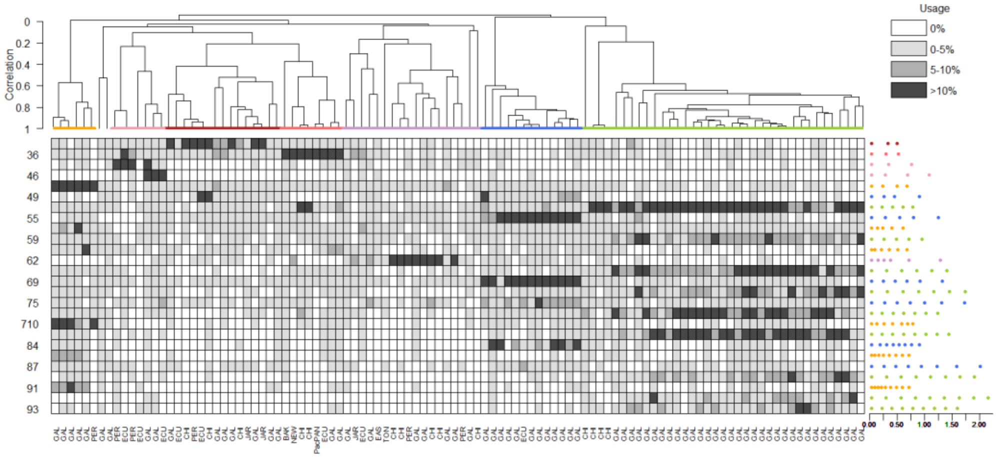
*Trial 6 (linkage=single)*



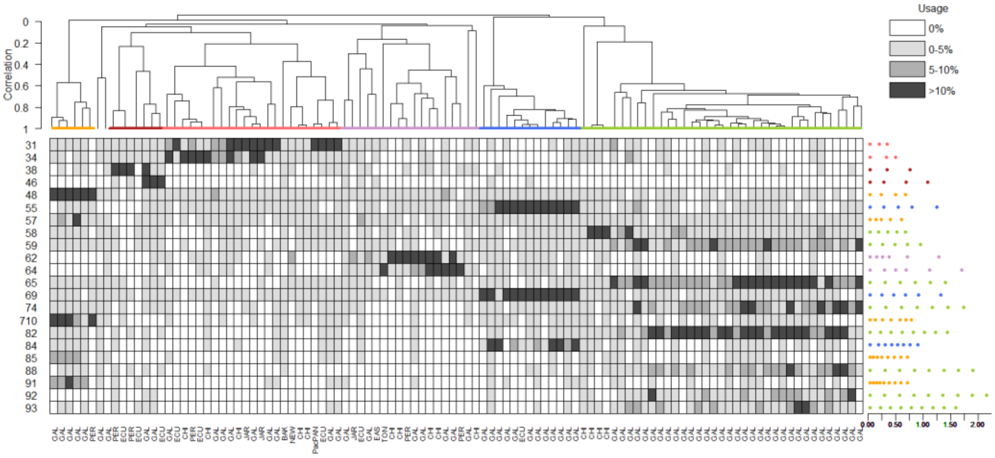
*Trial 7 (linkage=complete)*



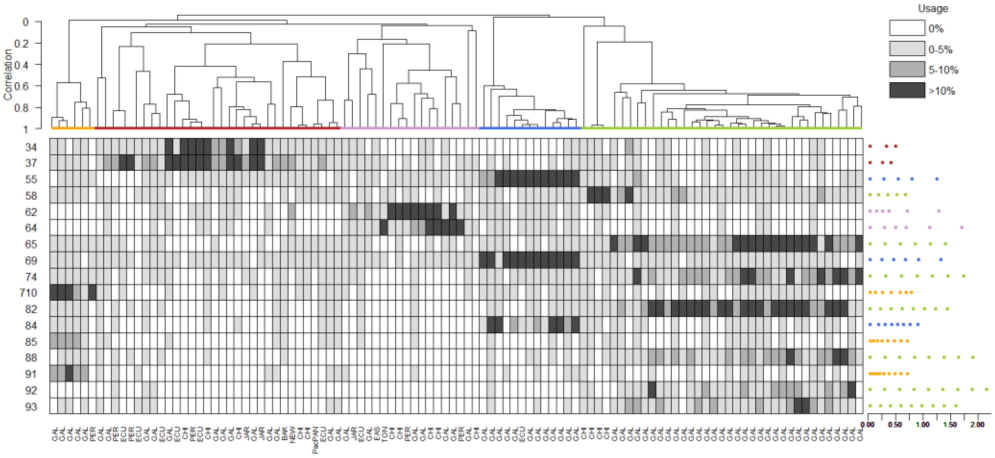
*Trial 8 (critfact=6)*



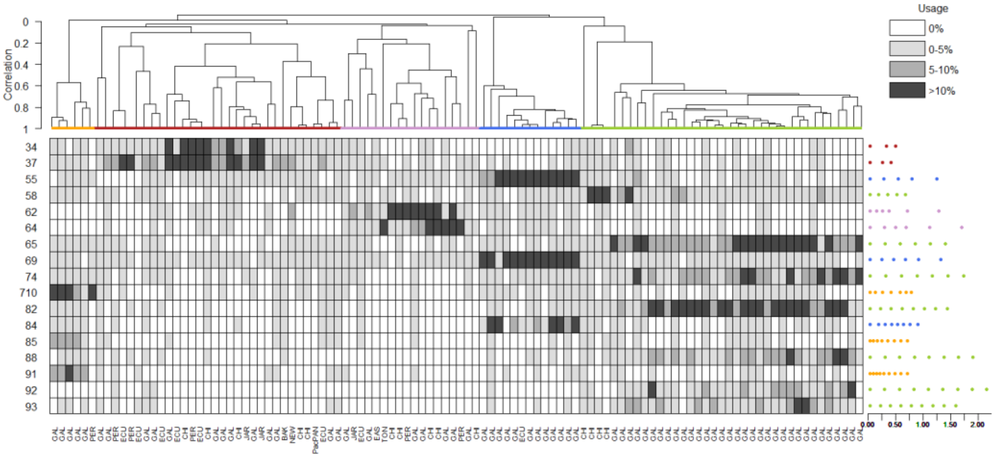
*Trial 9 (critfact=10)*



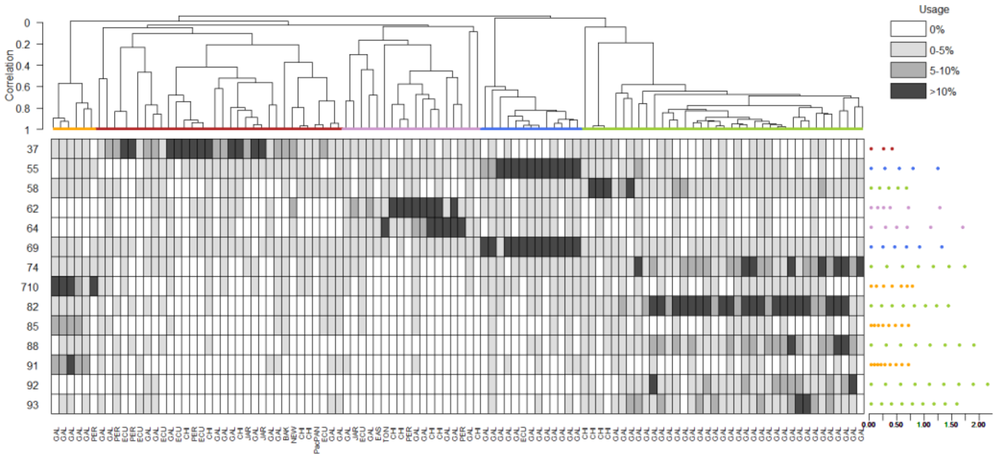
*Trial 10 (critfact=18)*



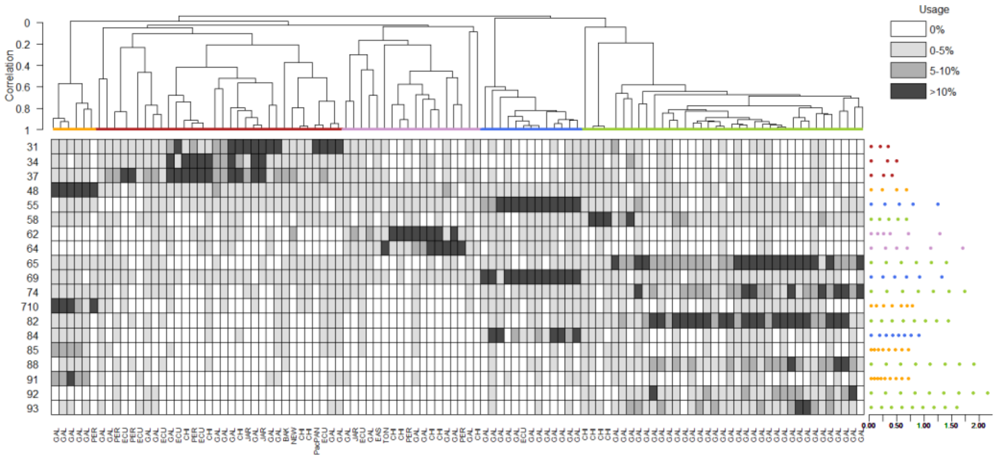
*Trial 11 (critfact=22)*



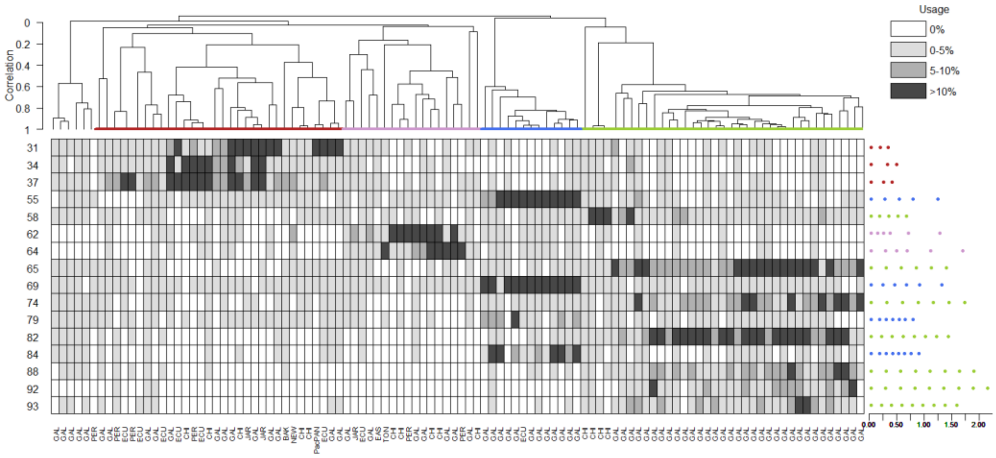
*Trial 12 (critfact=26)*



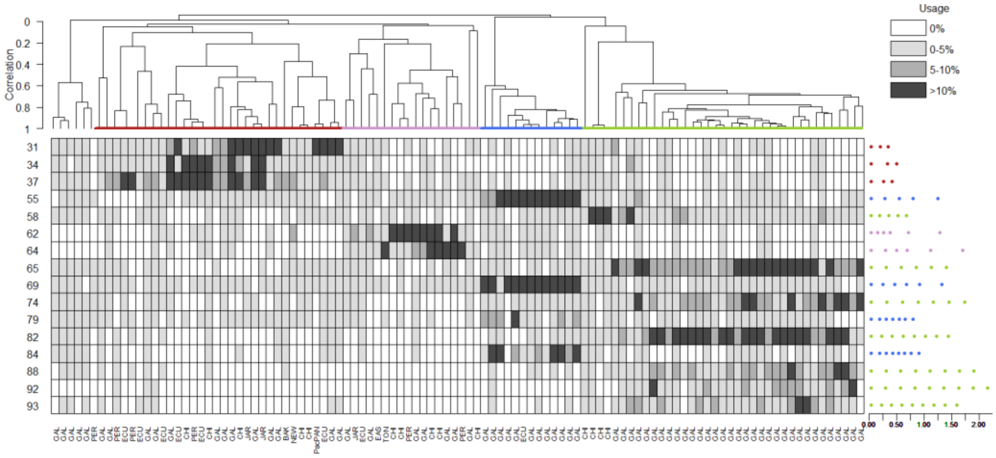
*Trial 13 (minrep=4)*



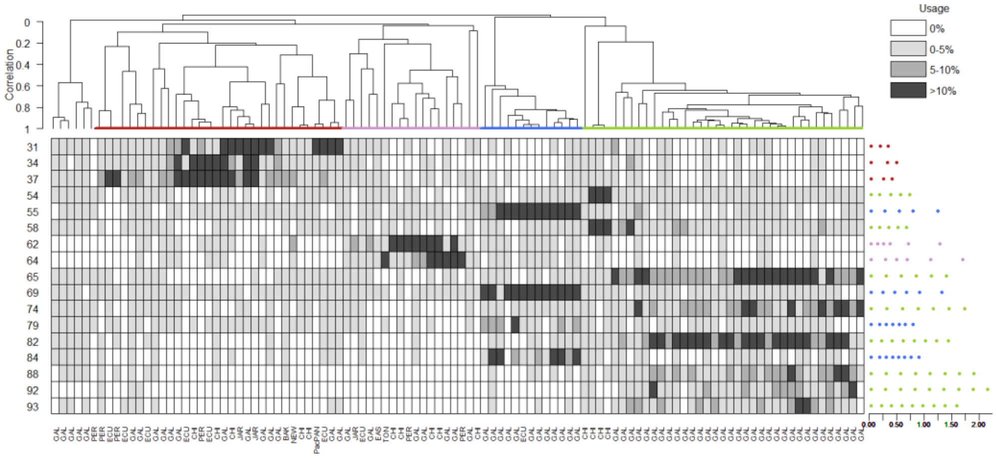
*Trial 14 (minrep=8)*



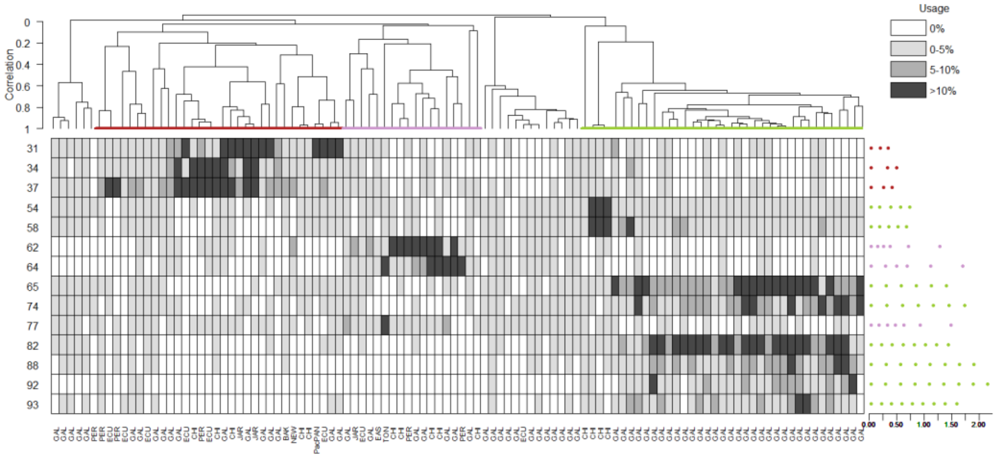
*Trial 15 (minrep=10)*



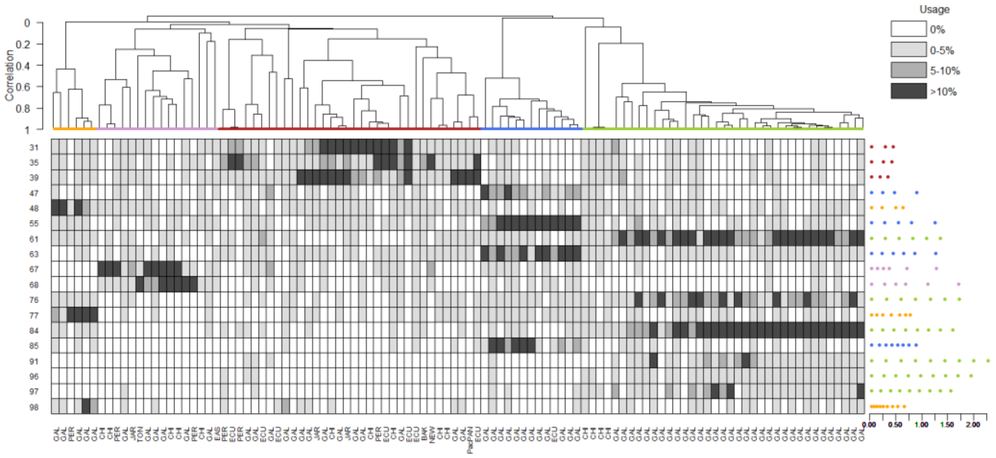
*Trial 16 (minrep=12)*



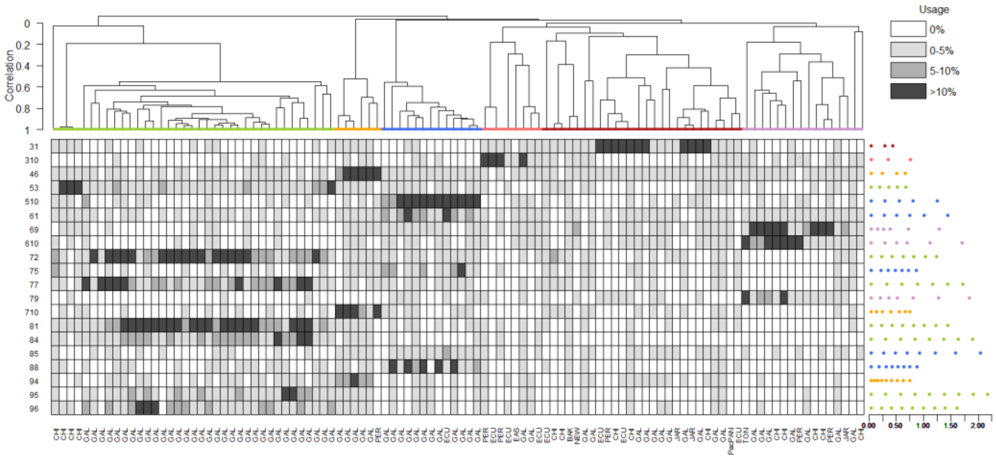
*Trial 17 (minrep=14)*



*Trial 18 (baseline run 2)*

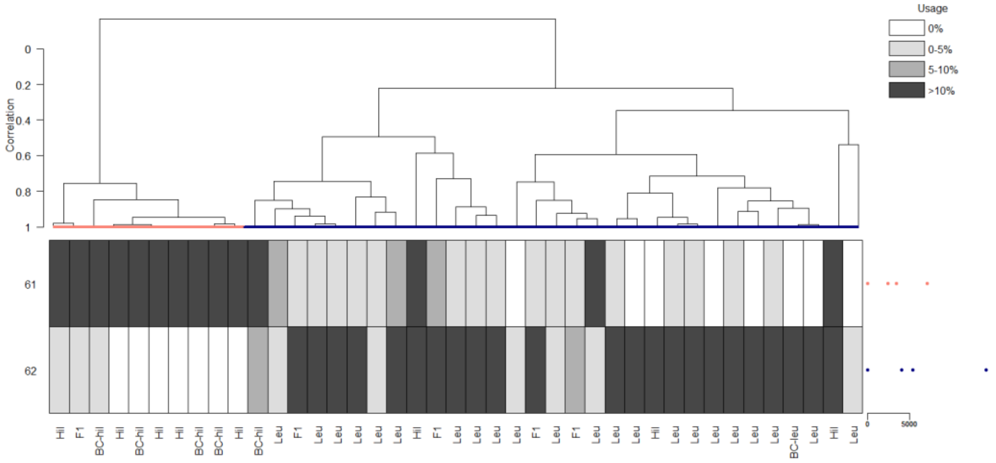


*Trial 19 (baseline run 3)*



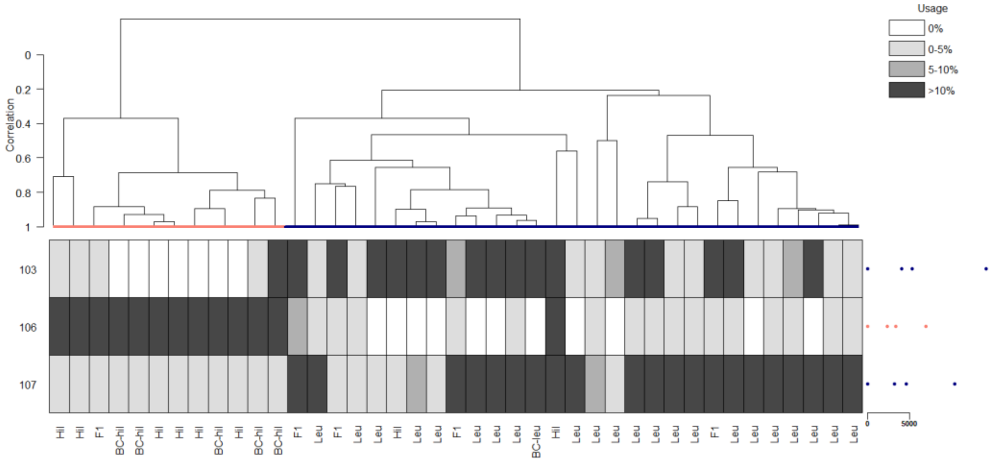
**B3.29:** *Data S3.3* – Wren trial results. The resultant dendrogram for each trial (n=19) is presented below, with trial numbers corresponding to those in Table S3.6. The varied option/parameter is listed after each trial number. All other options/parameters were kept at the default values (Table 3.2). For detailed figure description, see Figure 3.3. Acoustic clade colors match those used in Figure 3.3. Numeric call type codes (i.e. heat map row labels) are consistent for the baseline dendrogram and the trials varying the hierarchical clustering linkage strategy, *critfact*, and *minrep*, but not necessarily for the trials varying the ECM algorithm initialization strategy, the information criterion used during call classification, or the default value reruns.

*Trial 1 (initialization=random.post)*



*Trial 2 (initialization=random.clas):* No identity calls or clades delineated

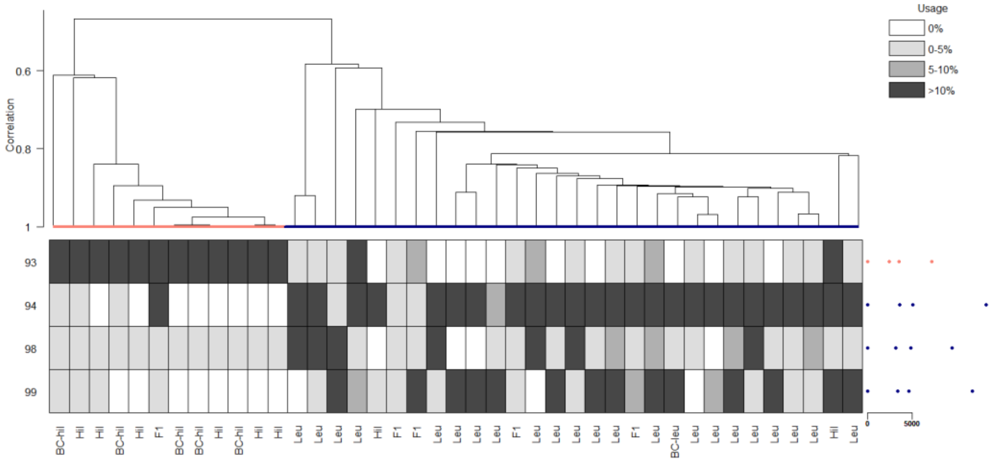
*Trial 3 (information criterion=AIC)*



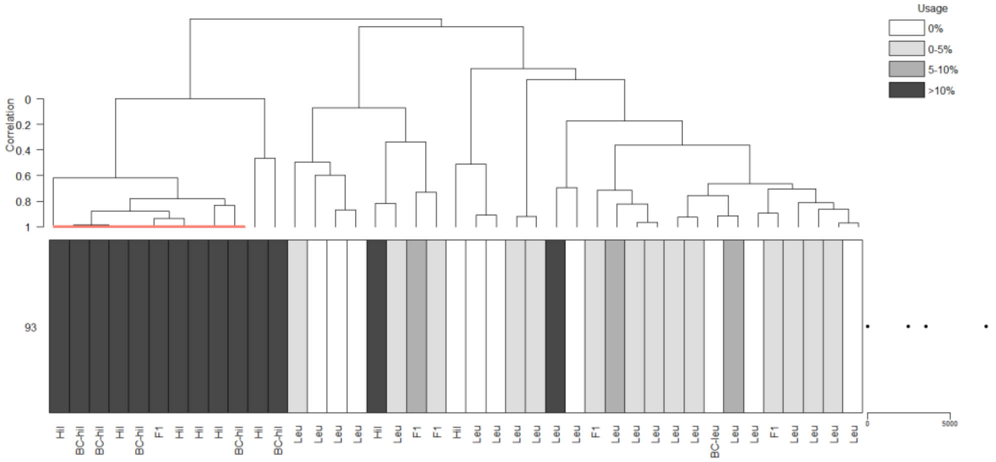
*Trial 4 (information criterion=BIC):* No identity calls or clades delineated

*Trial 5 (information criterion=ICL):* No identity calls or clades delineated

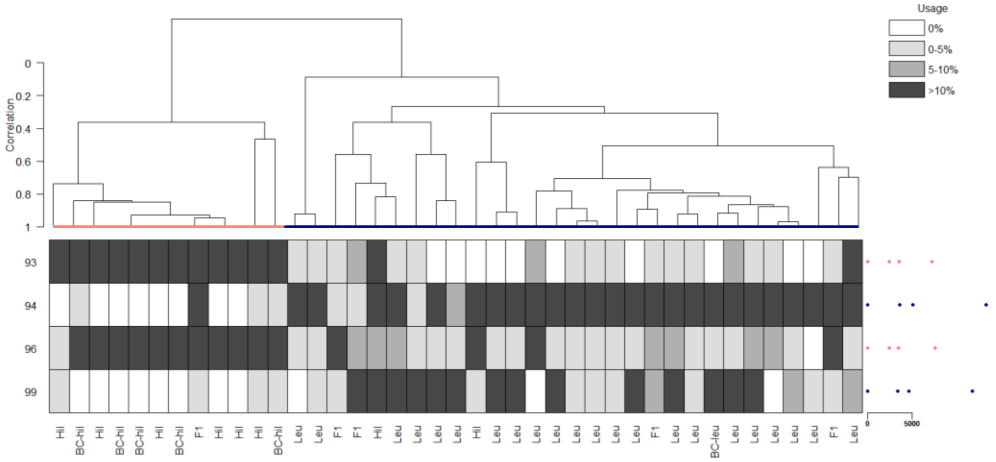
*Trial 6 (linkage=single)*



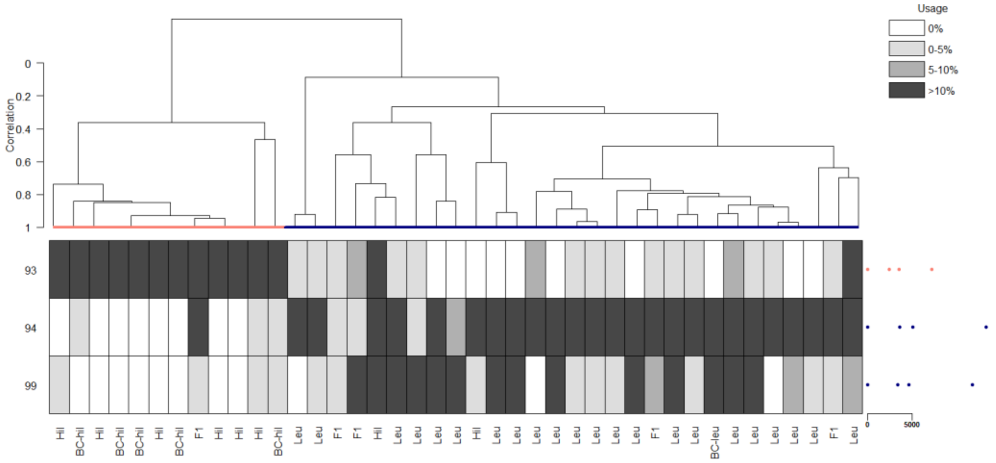
*Trial 7 (linkage=complete)*



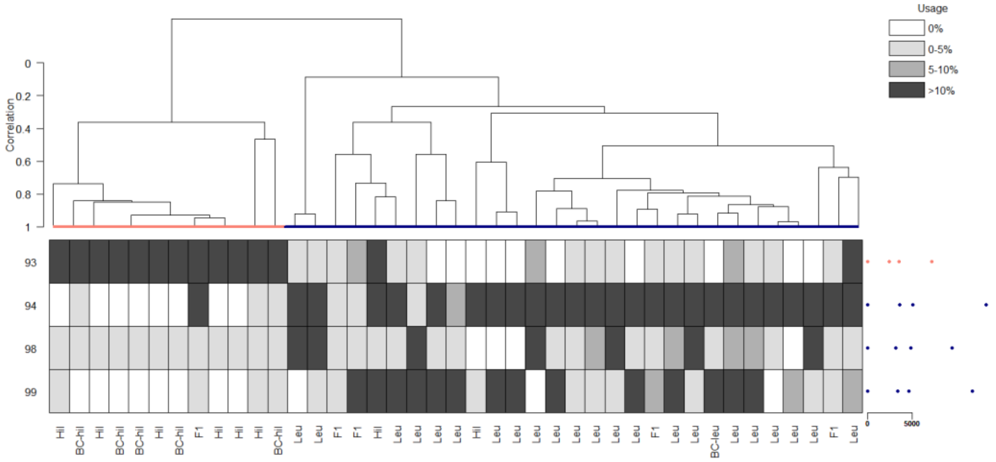
*Trial 8 (critfact=6)*



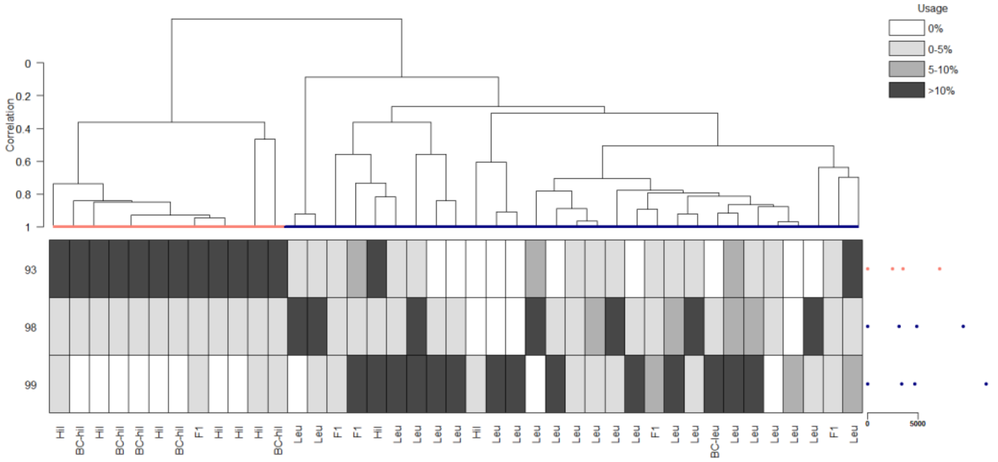
*Trial 9 (critfact=10)*



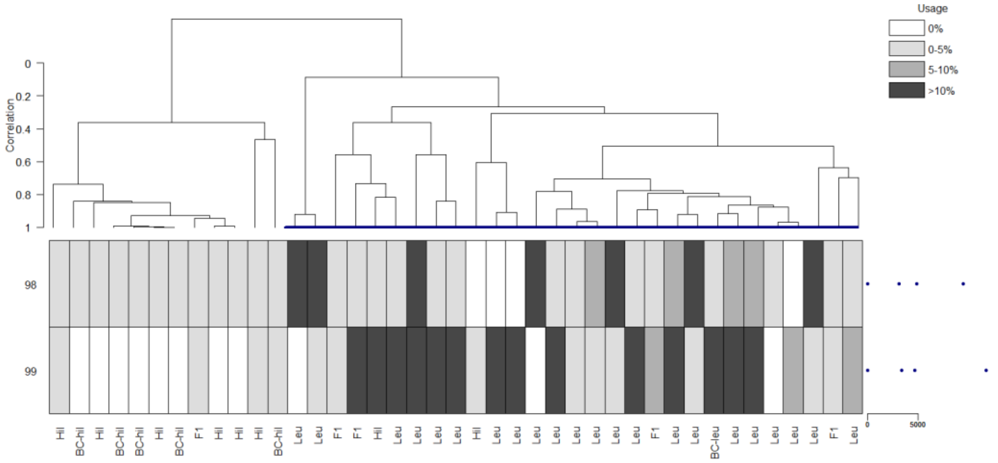
*Trial 10 (critfact=18)*



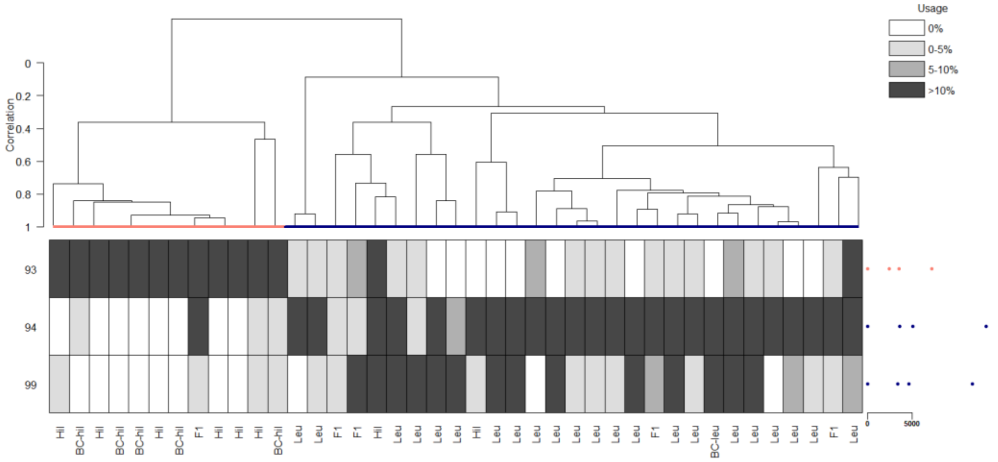
*Trial 11 (critfact=22)*



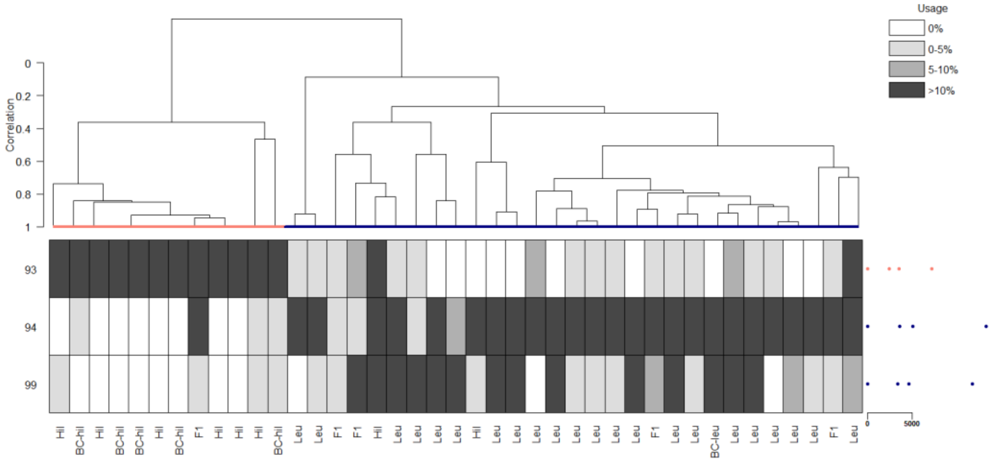
*Trial 12 (critfact=26)*



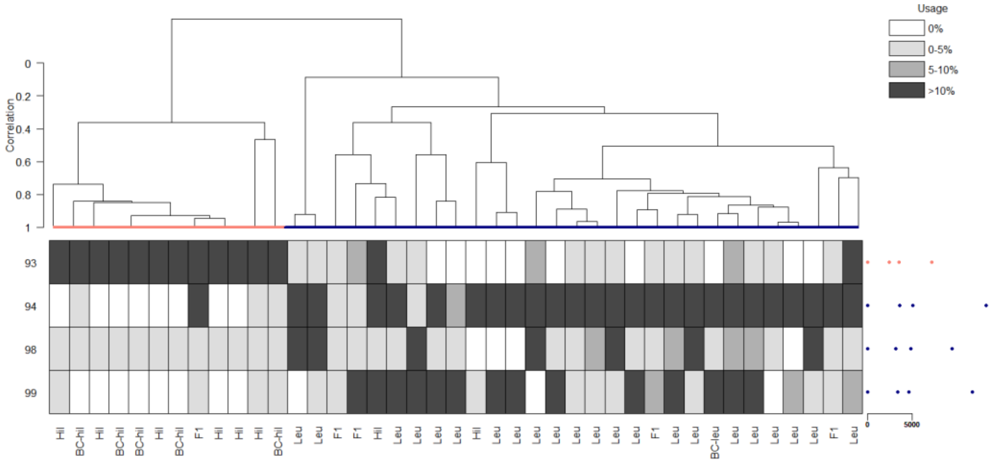
*Trial 13 (minrep=3)*



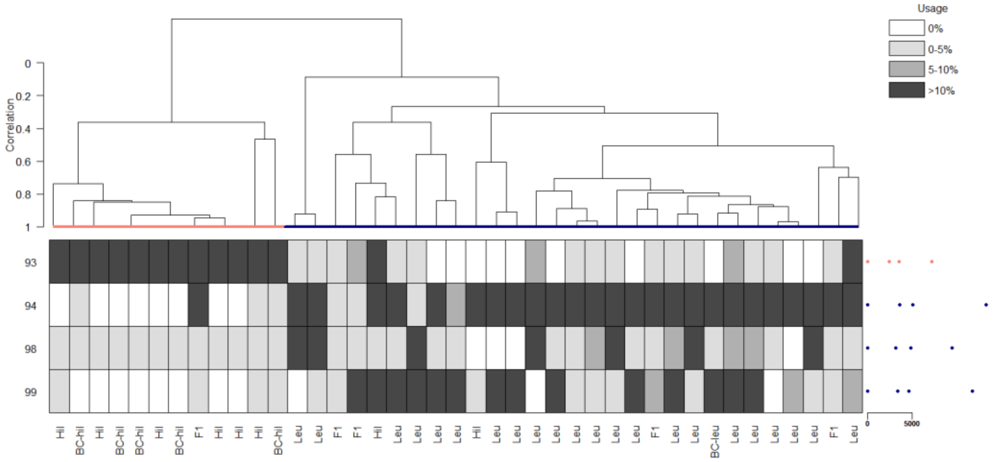
*Trial 14 (minrep=5)*



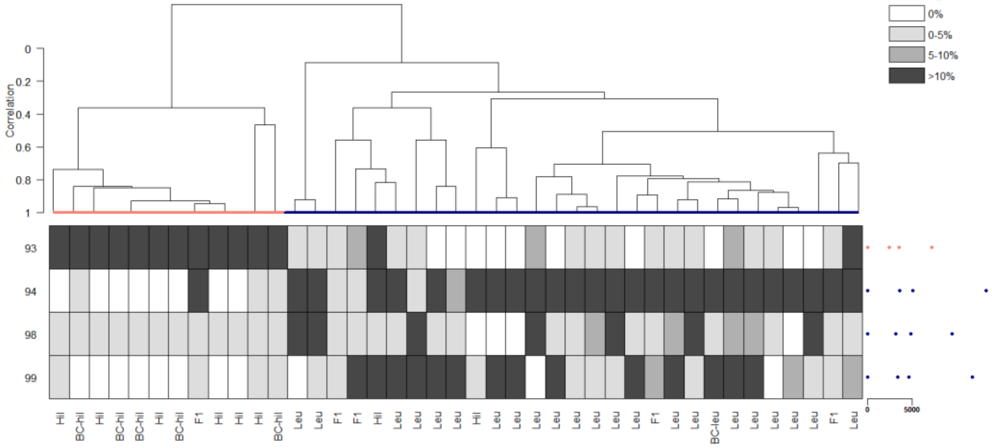
*Trial 15 (minrep=7)*



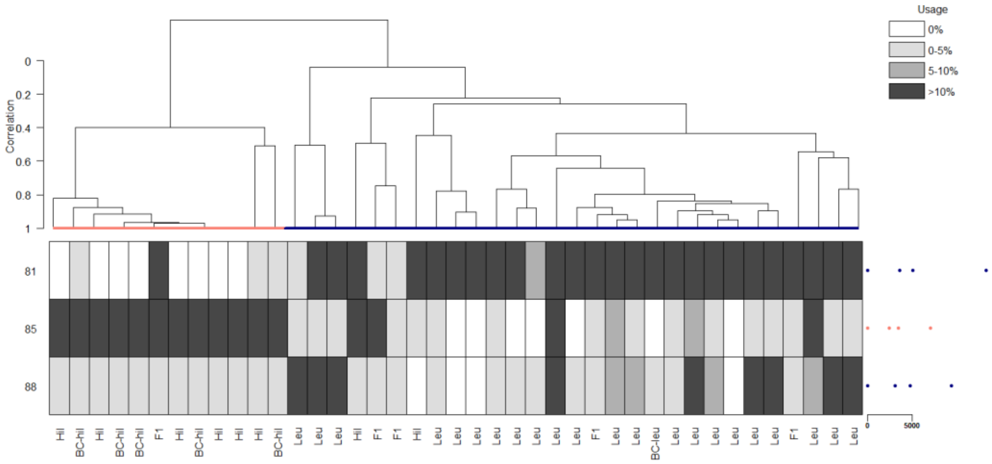
*Trial 16 (minrep=9)*



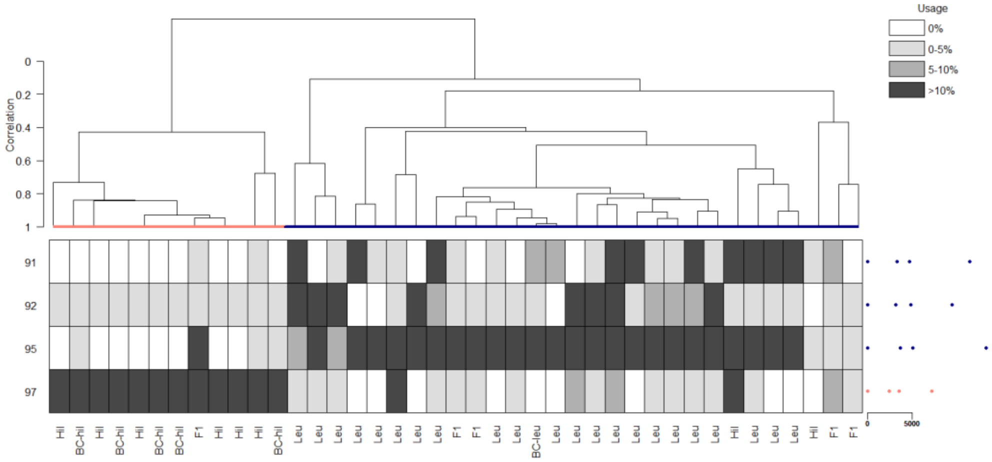
*Trial 17 (minrep=11)*



*Trial 18 (baseline run 2)*

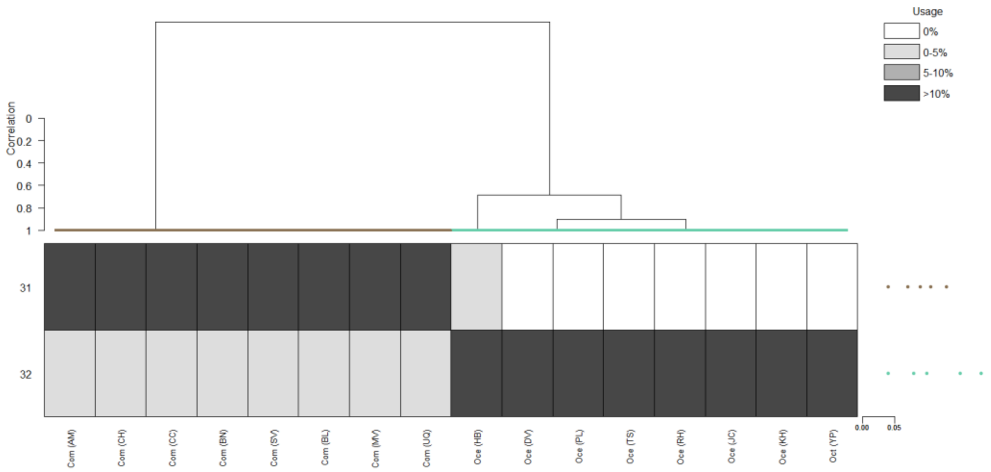


*Trial 19 (baseline run 3)*

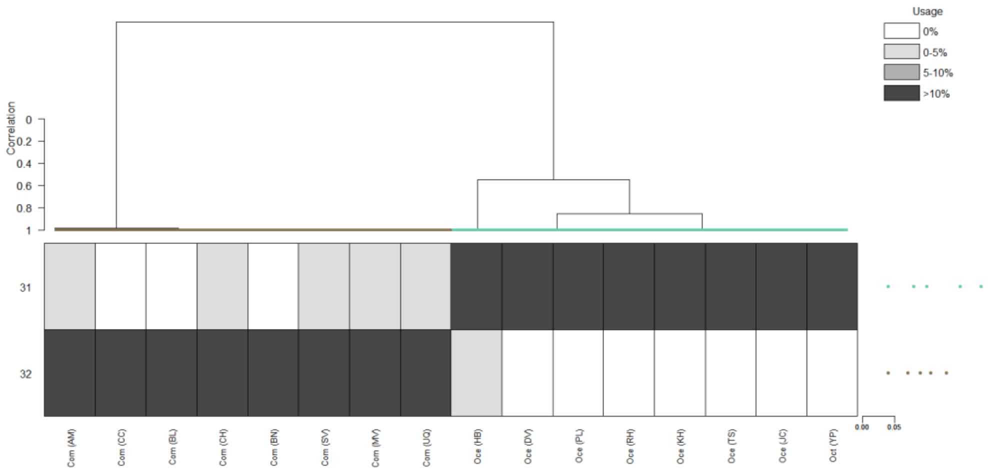


**B3.30:** *Data S3.4* – Cricket trial results. The resultant dendrogram for each trial (n=19) is presented below, with trial numbers corresponding to those in Table S3.7. The varied option/parameter is listed after each trial number. All other options/parameters were kept at the default values (Table 3.2). For detailed figure description, see Figure 3.4. Acoustic clade colors match those used in Figure 3.4. Numeric call type codes (i.e. heat map row labels) are consistent for the baseline dendrogram and the trials varying the hierarchical clustering linkage strategy, *critfact*, and *minrep*, but not necessarily for the trials varying the ECM algorithm initialization strategy, the information criterion used during call classification, or the default value reruns.

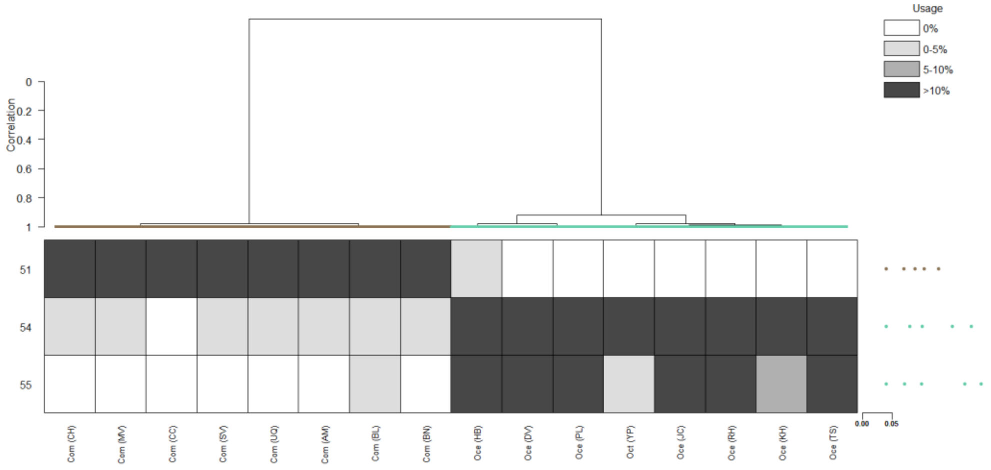
*Trial 1 (initialization=random.post)*



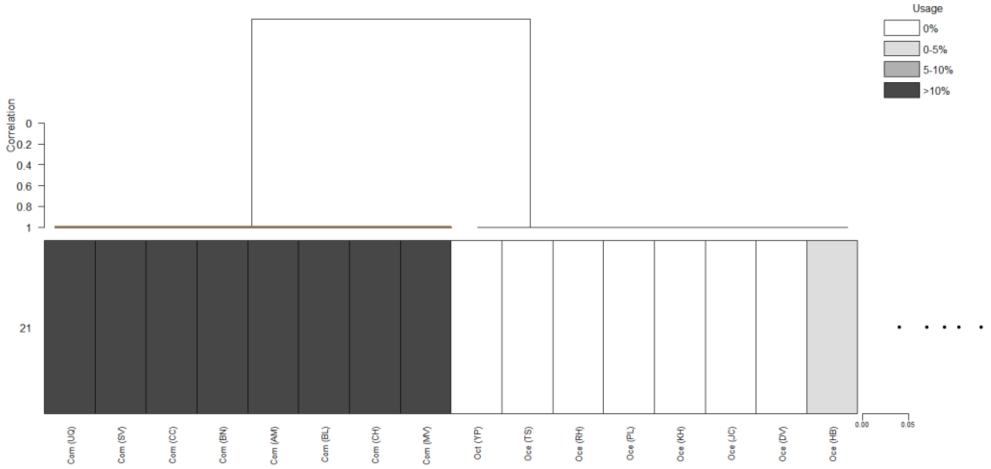
*Trial 2 (initialization=random.clas)*



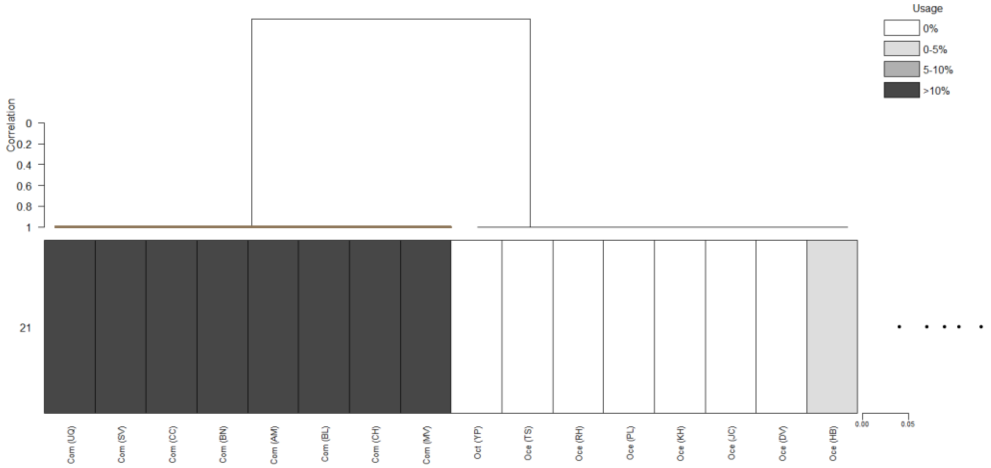
*Trial 3 (information criterion=AIC)*



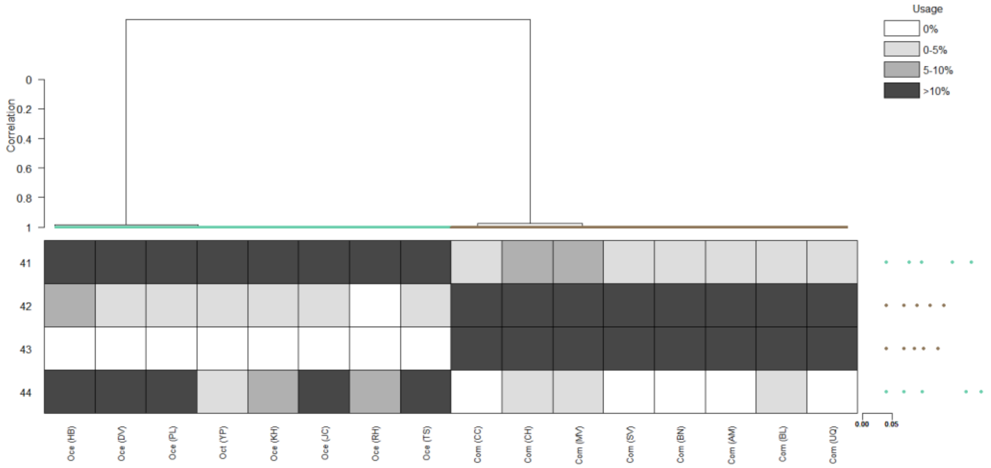
*Trial 4 (information criterion=BIC)*



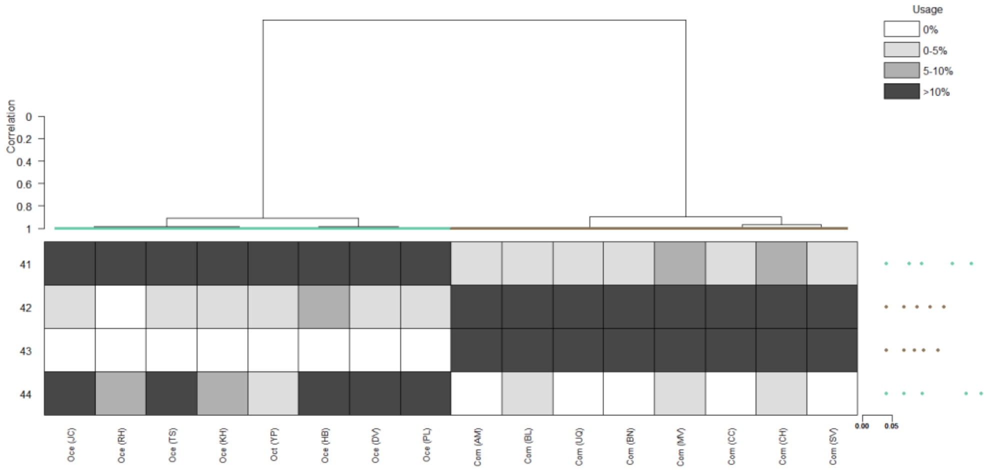
*Trial 5 (information criterion=ICL)*



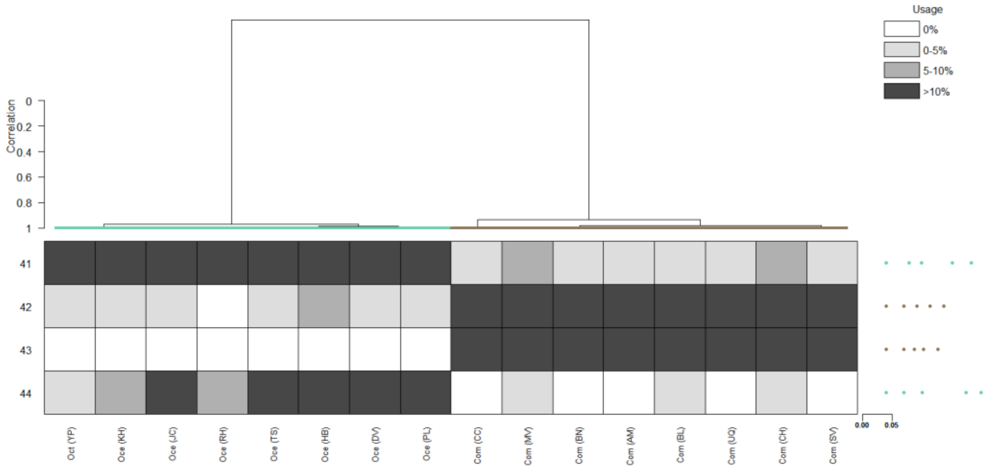
### *Trial 6 (linkage=single)*



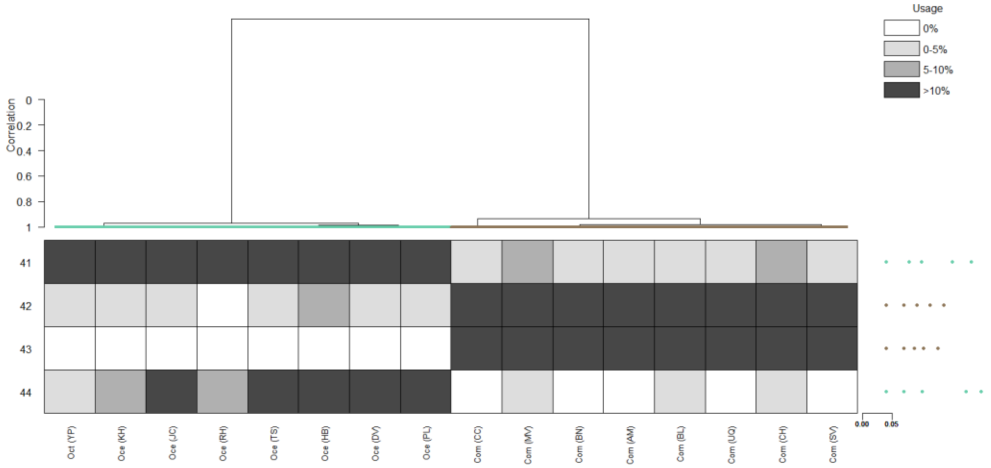
### *Trial 7 (linkage=complete)*



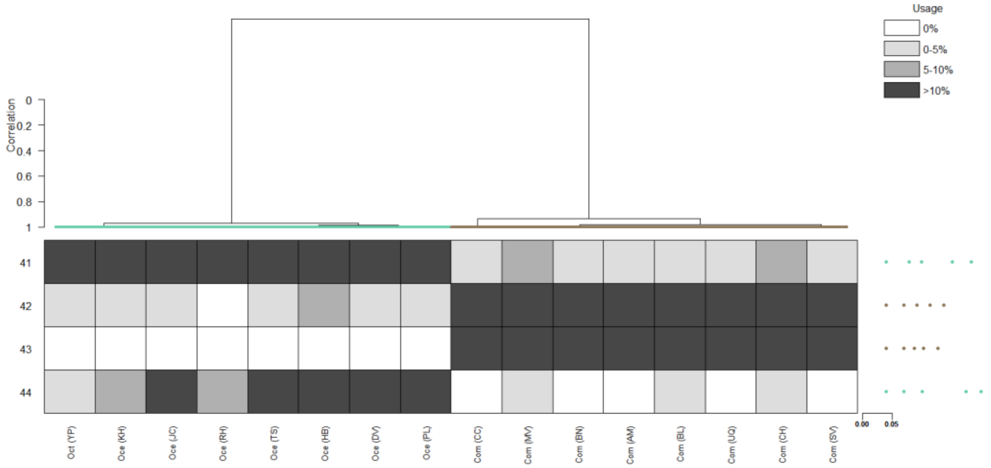
### *Trial 8 (critfact=6)*



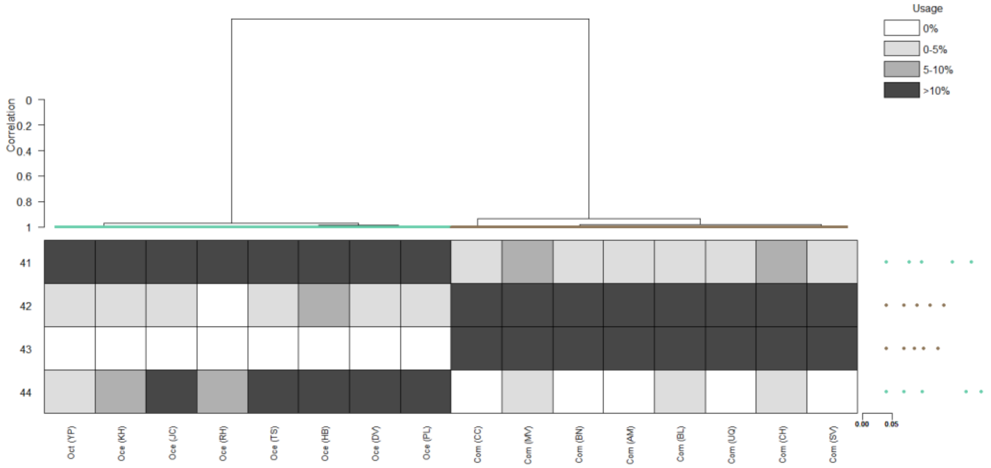
### *Trial 9 (critfact=10)*



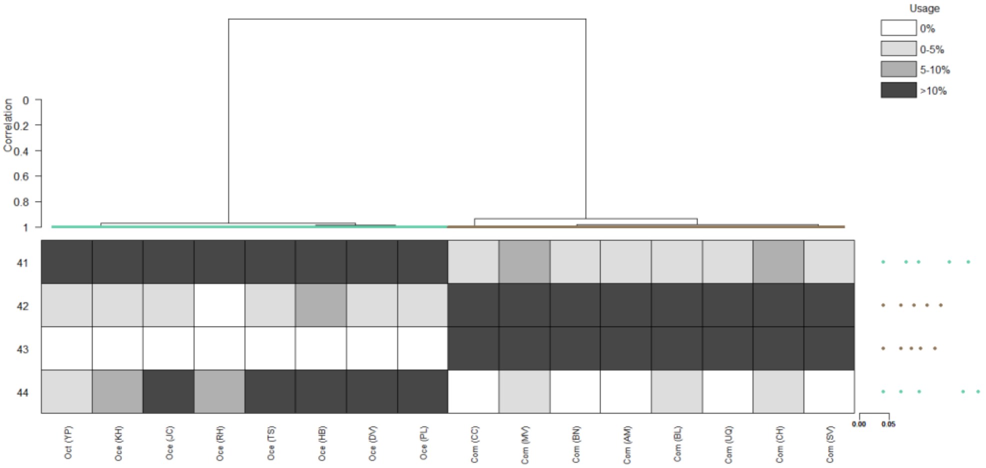
### *Trial 10 (critfact=18)*



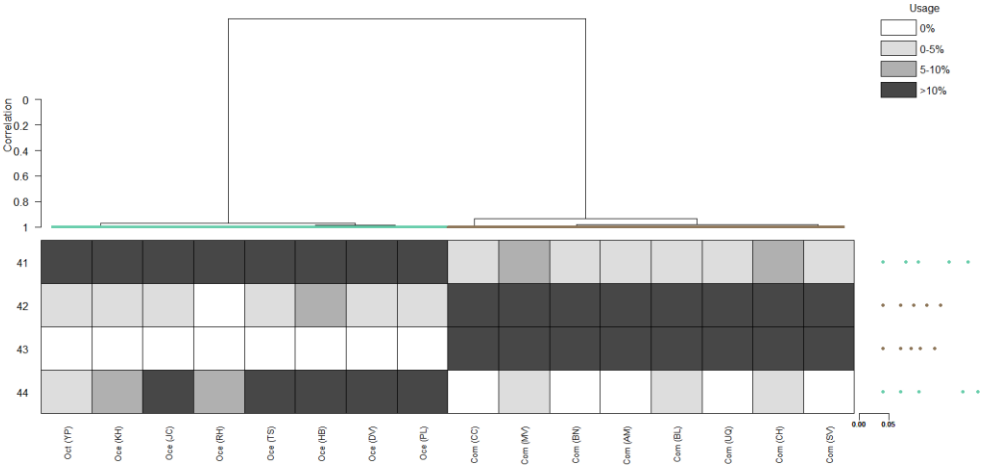
### *Trial 11 (critfact=22)*



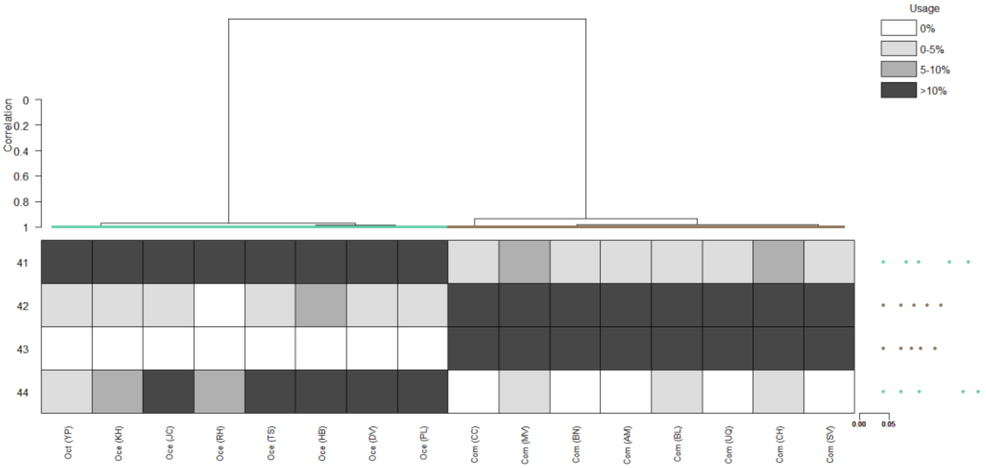
*Trial 12 (critfact=26)*



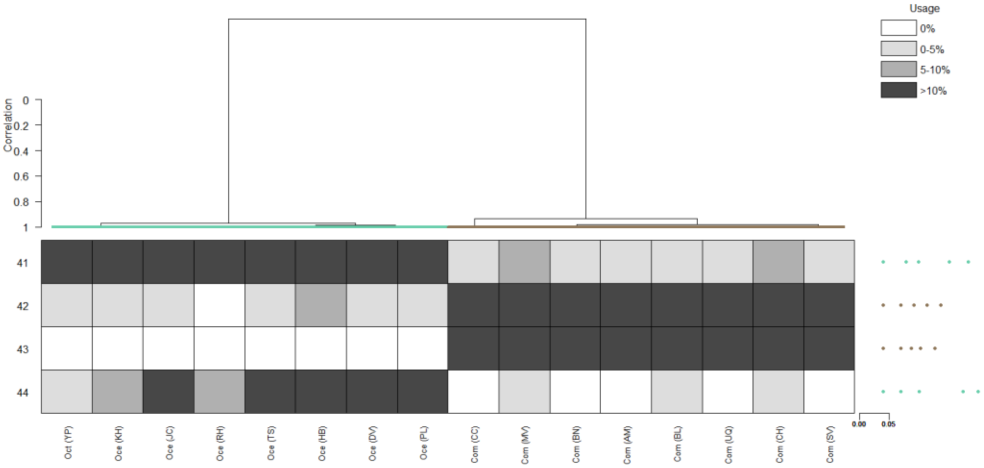
*Trial 13 (minrep=3)*



*Trial 14 (minrep=5)*



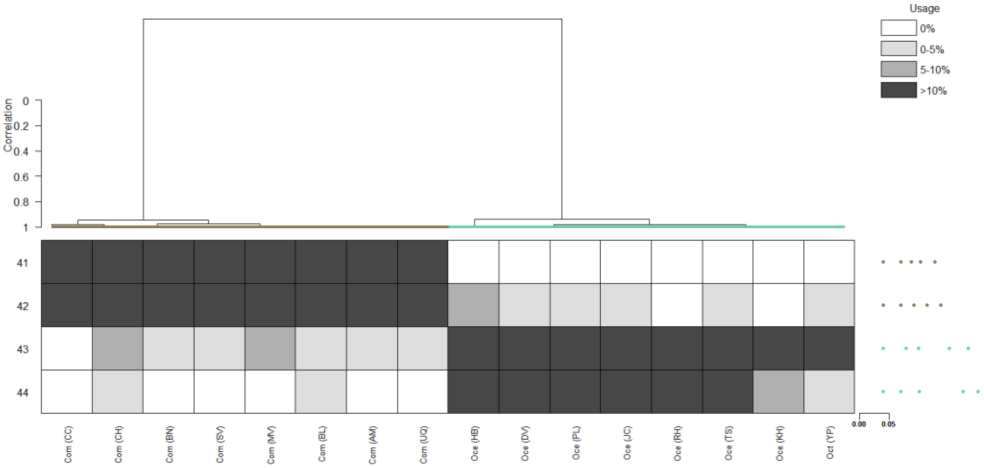
*Trial 15 (minrep=7)*



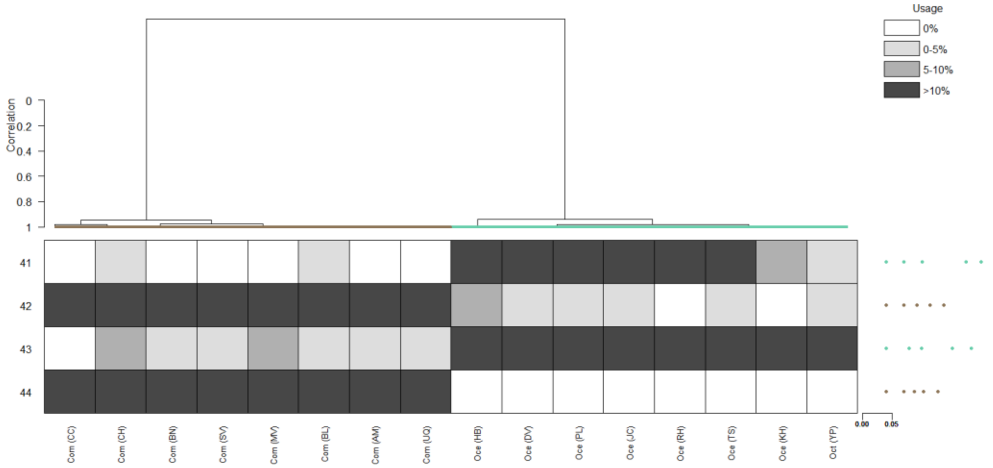
*Trial 16 (minrep=9):* No identity calls or clades delineated

*Trial 17 (minrep=11):* No identity calls or clades delineated

*Trial 18 (baseline run 2)*



### *Trial 19 (baseline run 3)*



# APPENDIX C – CHAPTER 4 SUPPLEMENT

**C4.1:** *Method S4.1* – Parameter setting deviations from Hersh et al. (2021)

We fitted 2:15 (instead of 2:10) mixture components to the data during call classification to account for the likely presence of additional coda types in this analysis compared to Hersh et al. (2021), given our expanded geographic scope and dataset size. The most computationally demanding step in IDcall is call classification, in which every combination of the set number of mixture components and each of a family of 14 models (Punzo & McNicholas, 2016) is fit to the data and compared using an information criterion. Using 2:15 mixture components and all 14 models results in 196 (i.e. 14x14) component/model combinations for each number of clicks (3–10 here), yielding 1,568 combinations total. To reduce the computational load, we performed preliminary data exploration using the 'mclust' R package (Scrucca et al., 2016) to determine if a subset of the 14 possible models consistently fit best (as determined by the Bayesian Information Criterion, BIC) the data for each coda click length. Models VVV, VVE, VEV, and VEE (see Punzo & McNicholas (2016) for model descriptions) were typically the top four models regardless of coda click length, with model VVV almost always best (Figure S4.1). Subsequently, we used those four models during the call classification stage of IDcall.

**C4.2:** *Method S4.2* – Geographic distance between repertoire calculates

We used the 'marmap' R package (Pante & Simon-Bouhet, 2013) to calculate the distance between repertoires. Bathymetric data for the Pacific Ocean were imported into R from the ETOPO1 dataset (Amante & Eakins, 2009) hosted on the National Oceanic and Atmospheric Administration server at 60-minute resolution. Shortest path calculations between repertoires were restricted to waters 1 km or deeper, as female sperm whales typically inhabit this water depth (Rice, 1989; Whitehead, 2003). The coarseness of 60-minute resolution resulted in some repertoires (n=24) located in grid cells with <1 km water depth. We shifted those repertoire positions incrementally by half a degree in a cardinal direction until the water depth reached ≥1 km. Most repertoires only needed to be shifted half a degree (n=18), but a few had to be shifted one degree (n=5). A single repertoire from NZL_S had to be shifted one degree southeast. Sixty-minute resolution corresponds to grid cells spanning ~111x111 km, meaning that two repertoires recorded within 111 km of each other are assigned a distance of 0 km. To achieve higher resolution in distance calculations for geographically close repertoires, we used the ‘geosphere’ R package (Hijmans et al., 2015) to calculate the great-circle distance for all repertoires within 200 km of each other. This distance matrix was merged with the other, such that the distance between repertoires separated by <200 km was calculated using the great-circle distance and the distance between repertoires separated by >200 km was calculated using bathymetric data with the 1 km depth minimum. These adjustments (jittering perfectly overlaid repertoires, shifting repertoires into deeper waters, etc.), coupled with sampling limitations (i.e. variable precision of localization data available for different regions; Table S4.1), meant that the between-repertoire distance calculations were approximate. However, the main distance distinctions of interest—very close vs. very far—were well preserved.

### **C4.3:** *Discussion S4.1* – Evidence in support of the 7-clan tree

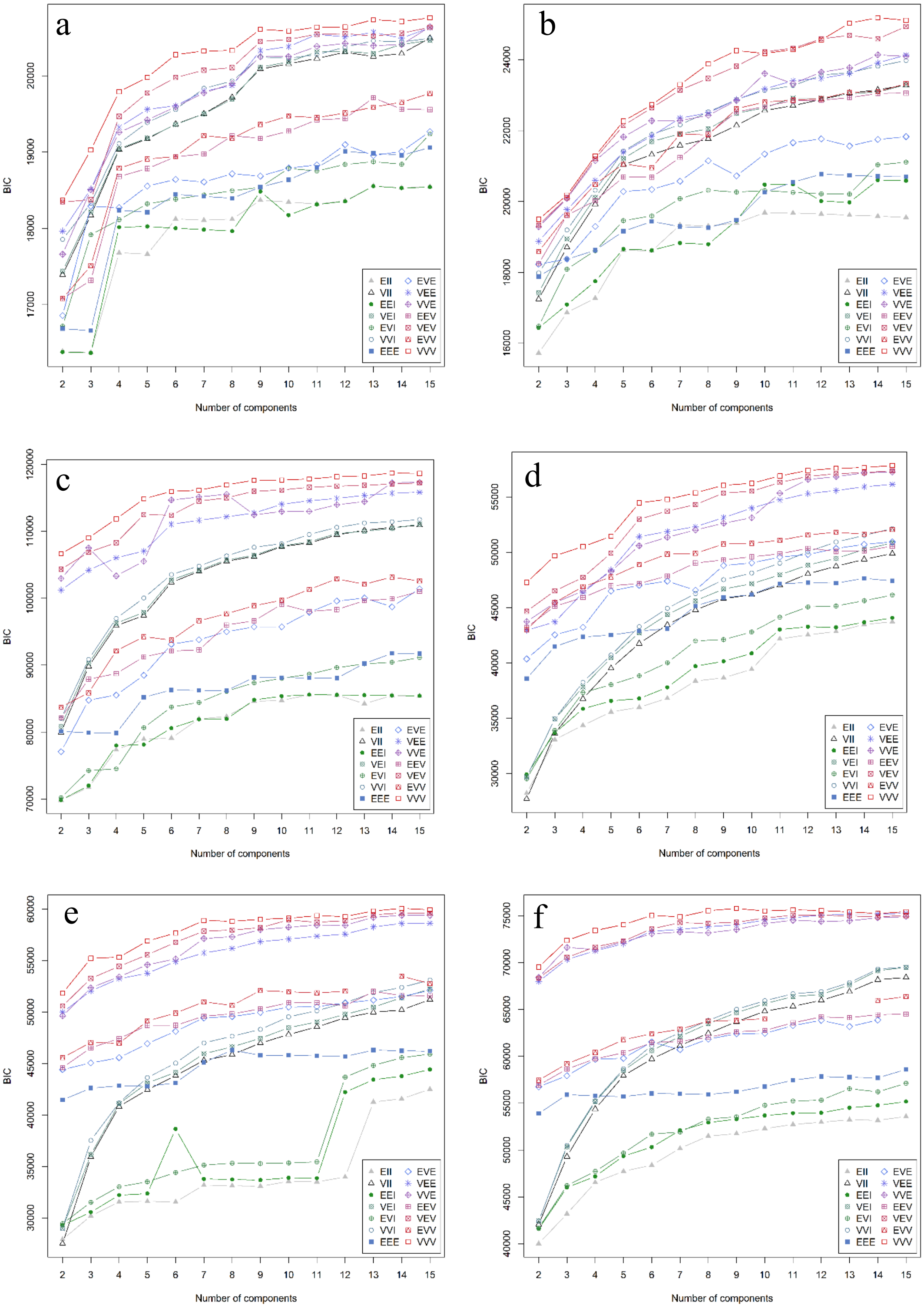
Photographic and acoustic work by Amano et al. (2014) found evidence for two clans off Japan, with whales recorded off the Ogasawara Islands sharing dialect similarities with the Short clan (which was first documented in the eastern tropical Pacific; Rendell & Whitehead, 2003b) and whales recorded off the Kumano coast belonging to a different clan. In the 5-clan scenario of the present analysis, the JPN_O and JPN_K repertoires are lumped into one large ‘Short’ clan (Figure S4.2). When we ran just the Japanese codas through IDcall, we consistently replicated the results from Amano et al. (2014), with the JPN_O and JPN_K repertoires clearly divided into two clans, even at extremely high values of *critfact* (Figure S4.19). This is promising, given several methodological differences between the two studies. For example, Amano et al. (2014) standardized ICIs by coda length and used all recorded 3–10 click codas, while we used absolute ICIs and required 25 codas per repertoire (this latter requirement reduced our coda sample size compared to theirs). This suggests that the JPN_O and JPN_K repertoires do indeed belong to different clans, and that any tree lumping them together (like the 5-clan tree; Figure S4.2) is likely incorrect (but see Figure S4.20 for a version of the distribution map with five clans).

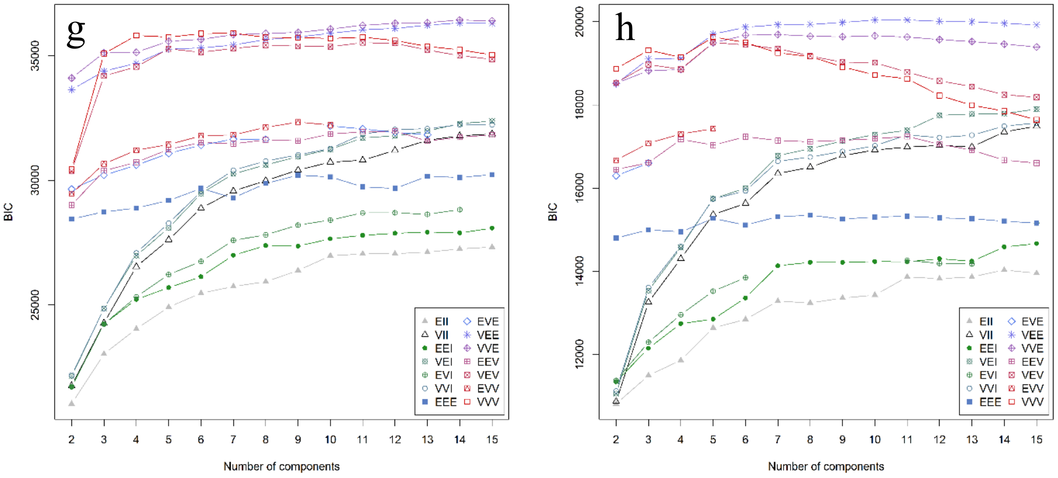
The 7- and 8-clan trees differ in that the 'Short' clan in the 7-clan tree (Figure 4.2, in red) is split into two clans in the 8-clan tree (Figure S4.5, in light blue and red), with repertoires primarily divided by longitude (i.e. West vs. East) (Figure S4.21). This longitudinal pattern fits well with within-clan spatial drift, rather than two fully diverged clans. This possibility is supported by the fact that the 'Short West' clan identity coda is rarely used (Figure S4.5), the 'Short East' clan identity coda (Figure S4.5) is the same as the 7-clan tree Short clan's identity coda (Figure 4.2), and both the Short West and the Short East clans primarily make three click codas (Figure S4.6). For these reasons, the 7-clan tree is the most compelling of the three scenarios and is the focus of subsequent analyses, but we emphasize that there is clearly more uncertainty in the clan structure of whales that predominantly make codas with fewer clicks.

### **C4.4:** *Discussion S4.2* – Evidence in support of the Short clan dialect as basal

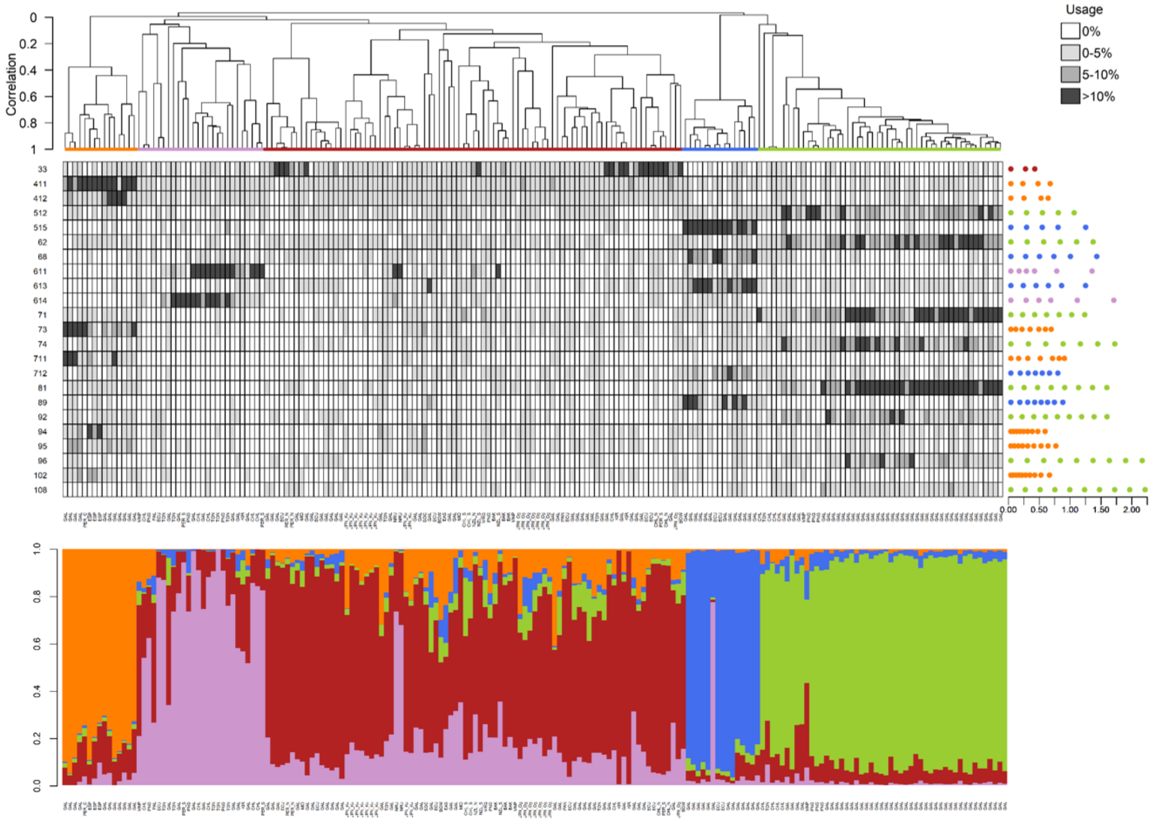
From a production standpoint, the Short clan's isochronous (i.e. regularly spaced), 3-click identity coda is one of the rhythmically simplest codas that can be made. Isochronous signals are prevalent in acoustic communication systems across taxa (Ravignani & Madison, 2017), which could indicate that they are evolutionarily basal throughout the animal kingdom. There is also a trend in the plots of minimum spatial overlap vs. acoustic similarity in the between-clans analysis, where the plots show a lopsided, inverted V shape (see panel 'a' in Figures 4.5/S4.11–S4.14/S4.16). Clans with low spatial overlap have low acoustic similarity (which fits with the pattern expected from drift); clans with intermediate spatial overlap have higher acoustic similarity; and clans with high spatial overlap have the lowest acoustic similarity. The net effect is a decrease in acoustic similarity with increasing clan overlap, but the increase in acoustic similarity at intermediate overlap is interesting. The clan pairs with intermediate overlap may have had a fairly recent common ancestral dialect, as evidenced by their high acoustic similarity. Closer examination of the clan pairs with intermediate overlap shows that the Short clan is typically one of the two clans in the pairs; our most parsimonious explanation is that the ancestral sperm whale dialect in the Pacific Ocean likely shared the most similarities with the modern Short clan dialect. Lastly, the Short clan is the primary clan detected at high latitudes (e.g. BOW, NZL_N, NZL_S, CHL_S; Figure 4.1), where the sperm whales are primarily males (Rice, 1989). Male sperm whales infrequently make codas and likely mate across clans (Rendell et al., 2005). We do not know if males adopt the clan dialect of the females they are with during courting or if they even make codas on these occasions, but the prevalence of Short clan detections at high latitudes suggests that males primarily make Short clan-style codas when they are not with females.

**C4.5:** *Figure S4.1* – Results of preliminary data exploration using mclust. Plots show the BIC value for each of 14 models (Punzo & McNicholas, 2016) when fitting 2 to 15 mixture components to 3–10-click codas (panels a–h). A higher BIC indicates a better fit.

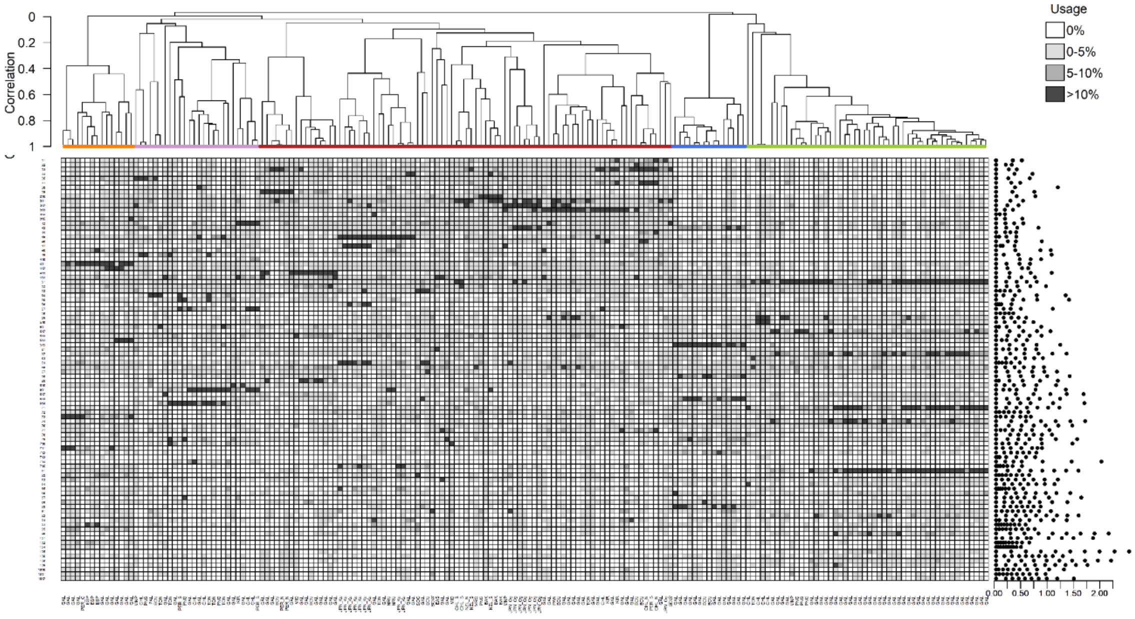




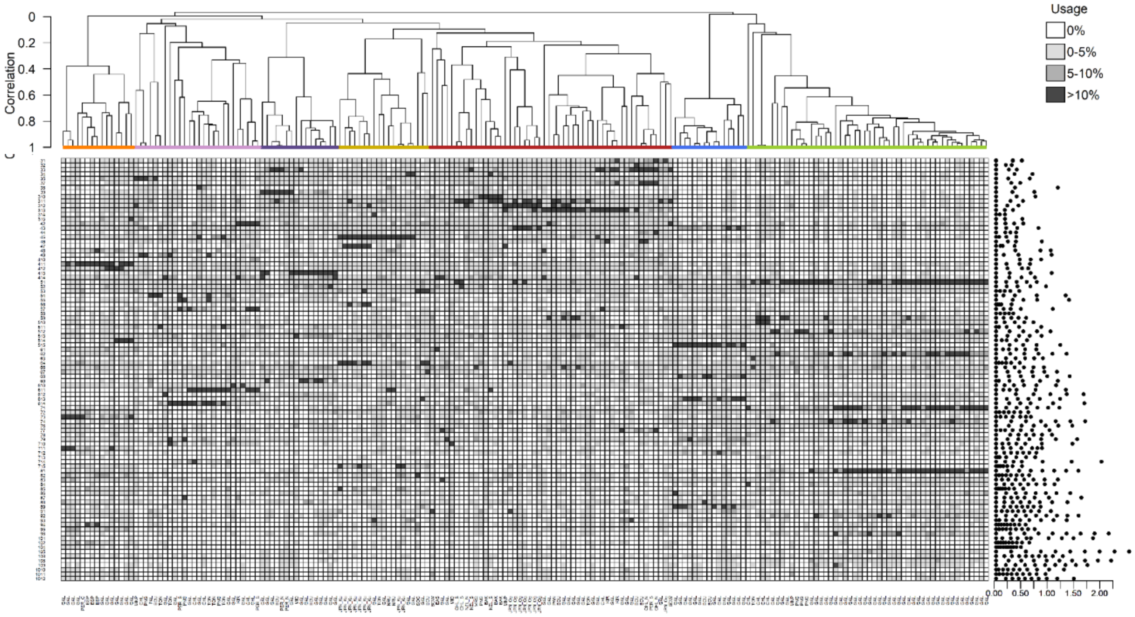
**C4.6:** *Figure S4.2* – Example 5-clan tree with identity coda types. This tree was constructed using *critfact*=12 and *minrep*=15 (Table S4.3). Colored identity clades correspond to five clans: Palindrome (orange), Four-Plus (pink), Short (red), Plus-One (blue), and Regular (green). See Figure 4.2 for additional details.



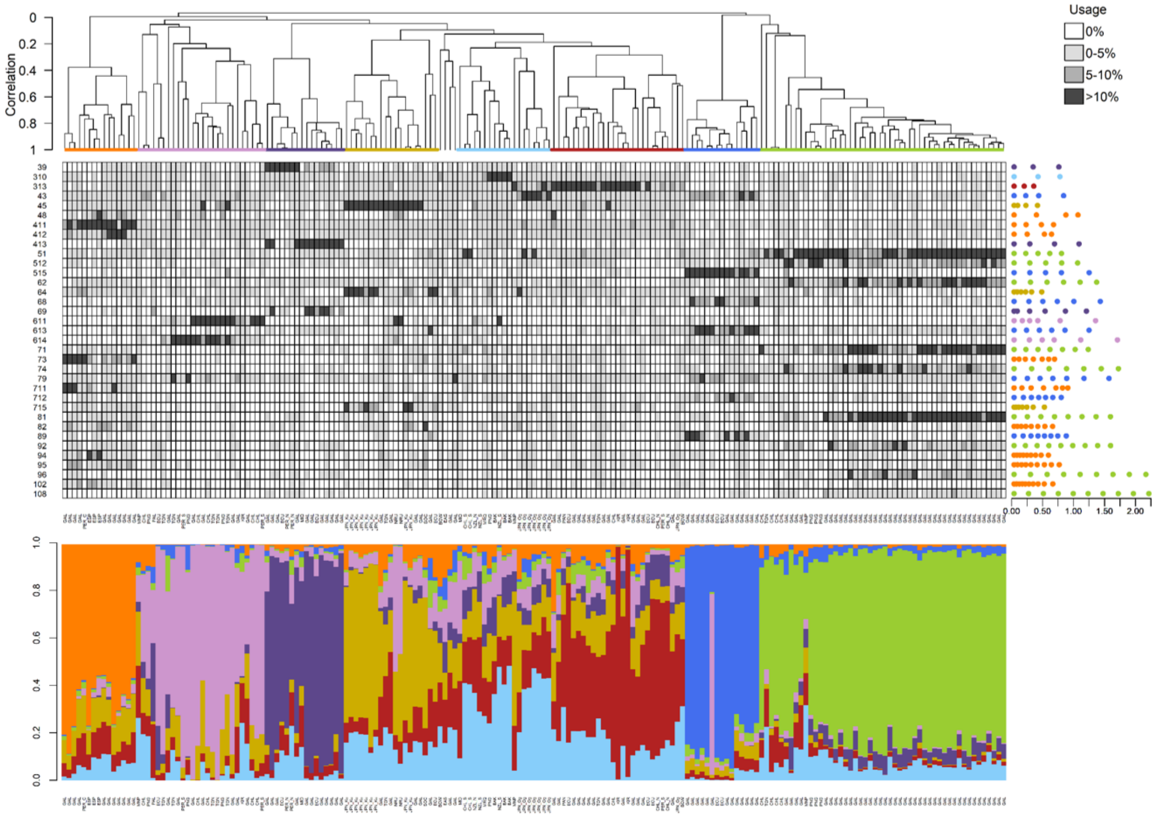
**C4.7:** *Figure S4.3* – Example 5-clan tree with all (identity and non-identity) coda types. This tree was constructed using *critfact*=12 and *minrep*=15 (Table S4.3). Colored identity clades correspond to five clans: Palindrome (orange), Four-Plus (pink), Short (red), Plus-One (blue), and Regular (green). See Figure 4.2 for additional details.



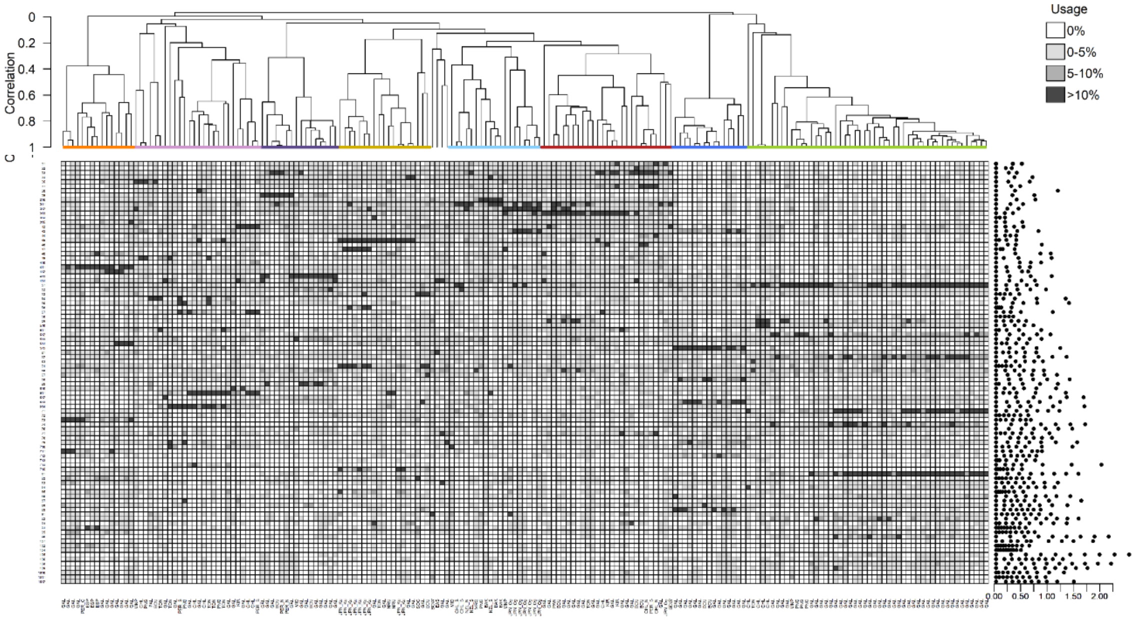
**C4.8:** *Figure S4.4* – Example 7-clan tree with all (identity and non-identity) coda types. This tree was constructed using *critfact*=5 and *minrep*=15 (Table S4.3). Colored identity clades correspond to seven clans: Palindrome (orange), Four-Plus (pink), Slow Increasing (purple), Rapid Increasing (gold), Short (red), Plus-One (blue), and Regular (green). See Figure 4.2 for additional details.



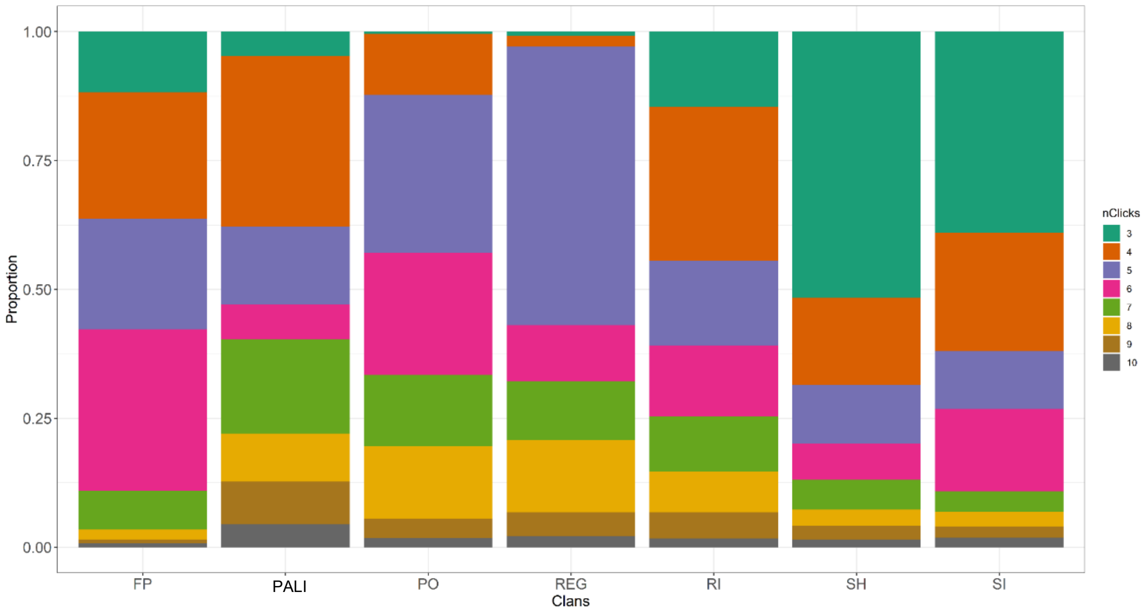
**C4.9:** *Figure S4.5* – Example 8-clan tree with identity coda types. This tree was constructed using *critfact*=6 and *minrep*=15 (Table S4.3). Colored identity clades correspond to eight clans: Palindrome (orange), Four-Plus (pink), Slow Increasing (purple), Rapid increasing (gold), Short West (light blue), Short East (red), Plus-One (dark blue), and Regular (green). Four repertoires (recorded off GAL, GAL, BOW, and EAS) are outliers/not assigned to a clan in this tree. See Figure 4.2 for additional details.



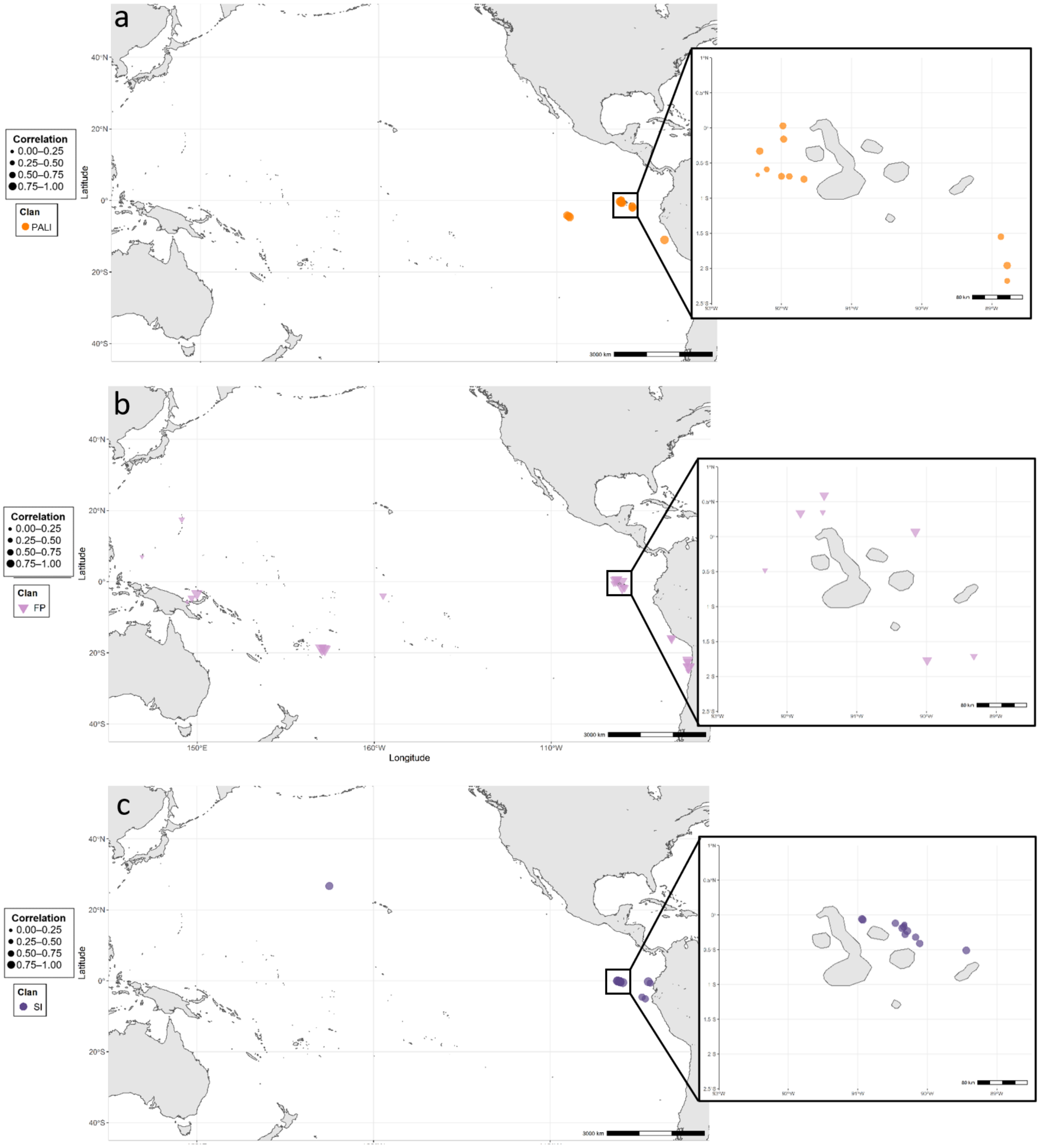
**C4.10:** *Figure S4.6* – Example 8-clan tree with all (identity and non-identity) coda types. This tree was constructed using *critfact*=6 and *minrep*=15 (Table S4.3). Colored identity clades correspond to eight clans: Palindrome (orange), Four-Plus (pink), Slow Increasing (purple), Rapid Increasing (gold), Short West (light blue), Short East (red), Plus-One (blue), and Regular (green). Four repertoires (recorded off GAL, GAL, BOW, and EAS) are outliers/not assigned to a clan in this tree. See Figure 4.2 for additional details.

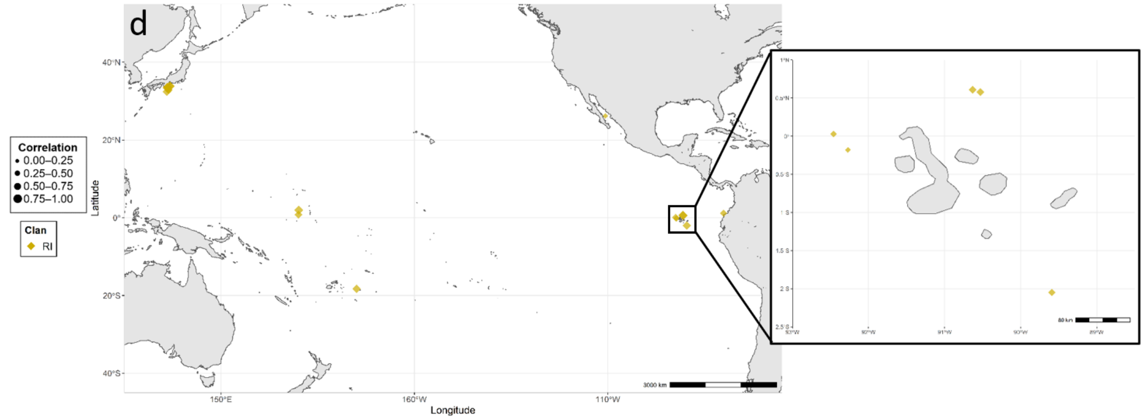


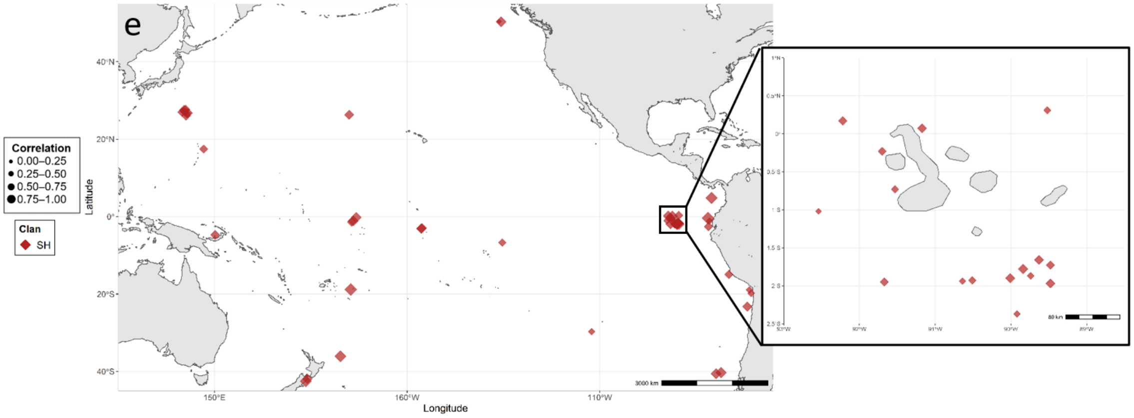
**C4.11:** *Figure S4.7* – Proportion of clan codas made up of different numbers of clicks. For clan abbreviations, see Figure 4.1.

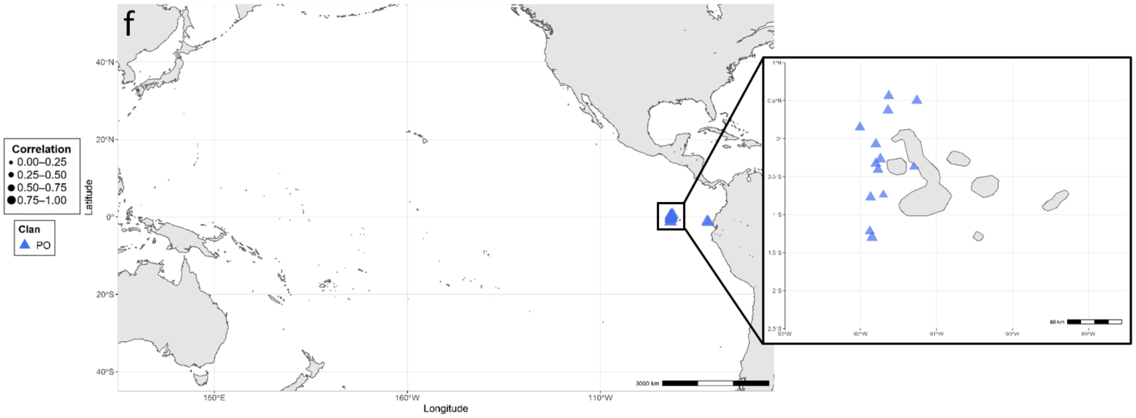


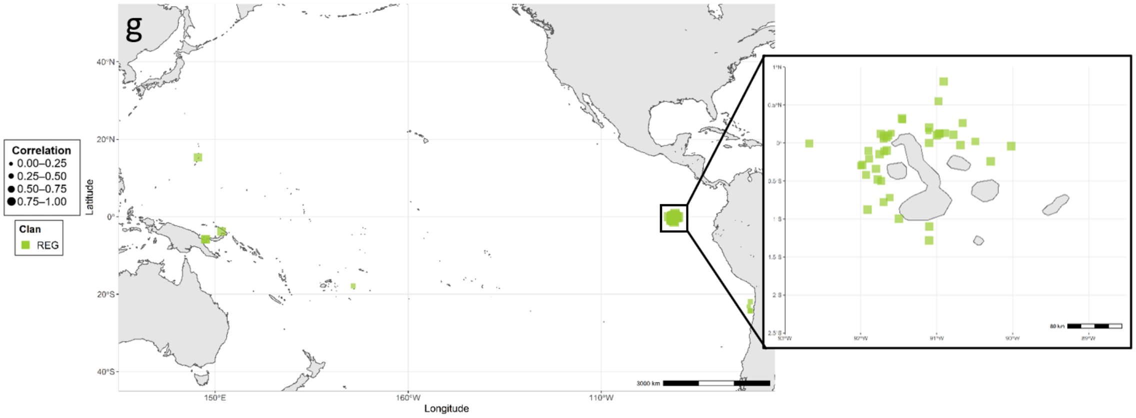
**C4.12:** *Figure S4.8* – Maps of individual clan distributions across the Pacific Ocean using the 7-clan tree clan designations. See Figure 4.1 for a composite map and for additional details. Panels correspond to the: (a) Palindrome; (b) Four-Plus; (c) Slow Increasing; (d) Rapid Increasing; (e) Short; (f) Plus-One; and (g) Regular clans.



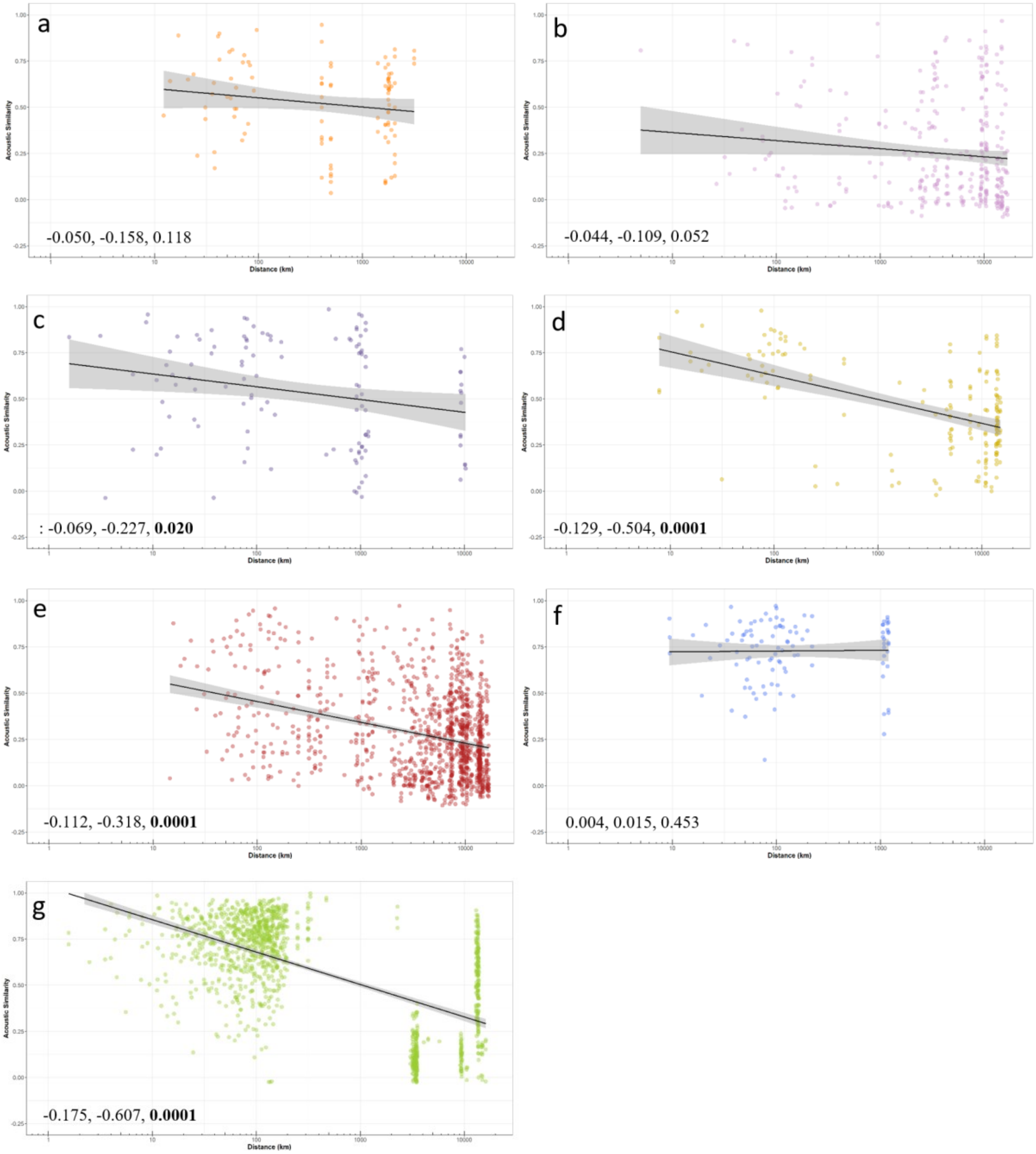

d
Correlation
0.00–0.25
0.25–0.50
0.50–0.75
0.75–1.00
Clan
RI
40°N
20°N
0°
20°S
40°S
Latitude
150°E
160°W
110°W
Longitude
3000 km
80 km


e
Correlation
0.00–0.25
0.25–0.50
0.50–0.75
0.75–1.00
Clan
SH
40°N
20°N
0°
20°S
40°S
Latitude
150°E
160°W
110°W
Longitude
3000 km
80 km

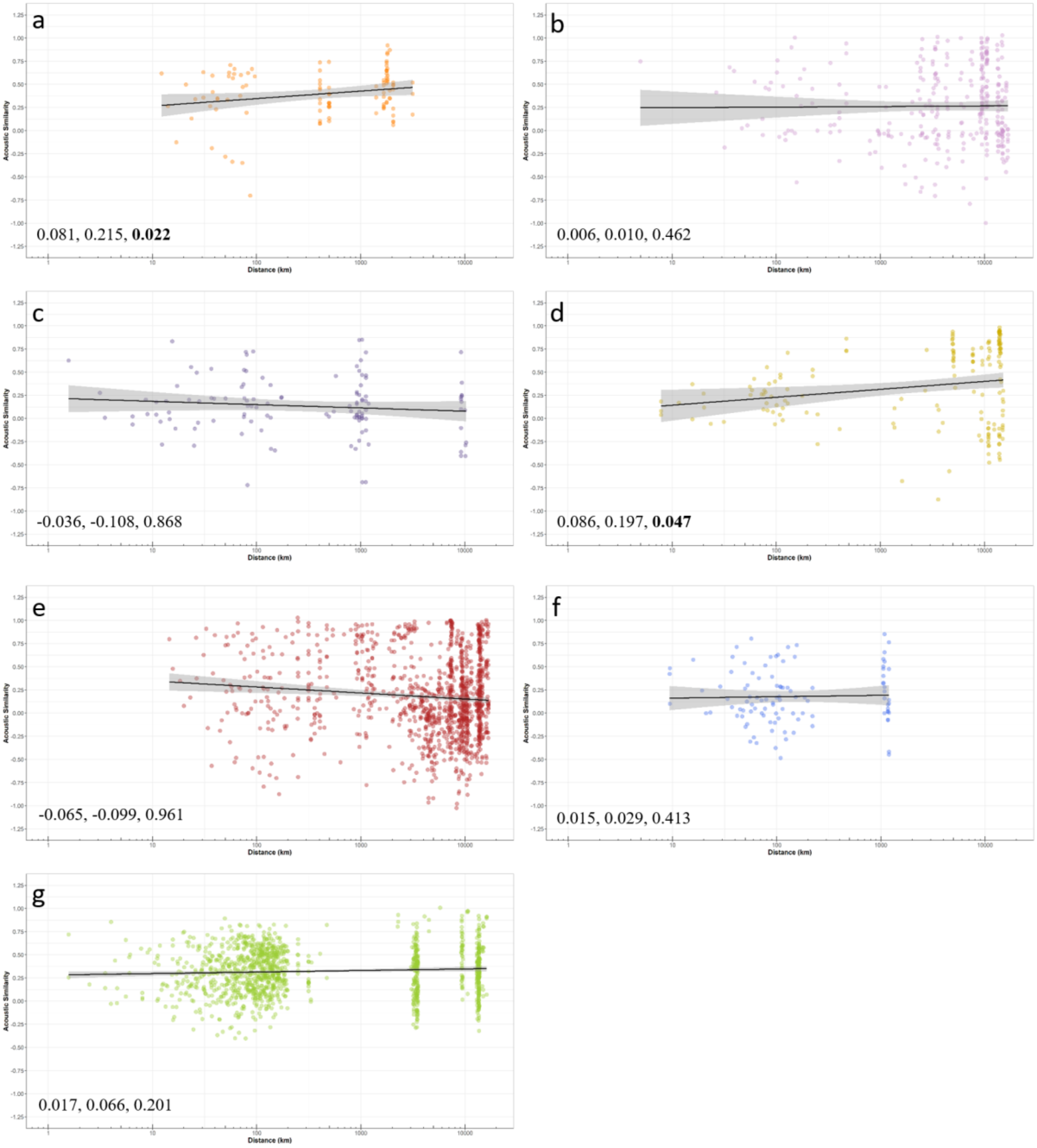

f
Correlation
0.00–0.25
0.25–0.50
0.50–0.75
0.75–1.00
Clan
PO
40°N
20°N
0°
20°S
40°S
Latitude
150°E
160°W
110°W
Longitude
3000 km
80 km


g
Correlation
0.00–0.25
0.25–0.50
0.50–0.75
0.75–1.00
Clan
REG
40°N
20°N
0°
20°S
40°S
Latitude
150°E
160°W
110°W
Longitude
3000 km
80 km

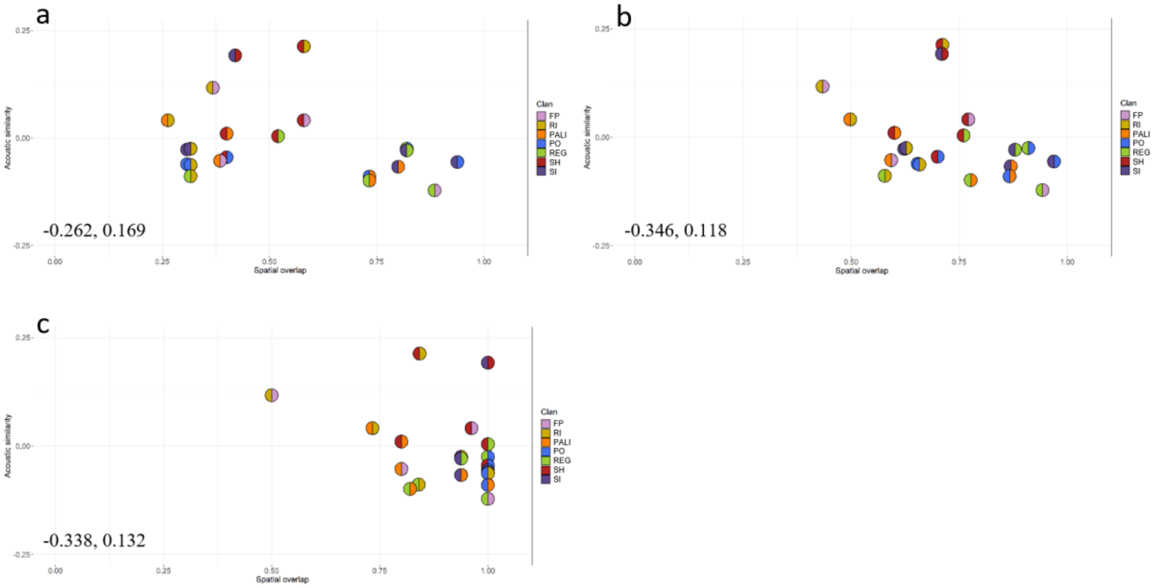
**C4.13:** *Figure S4.9* – Logged geographic distance vs. $\text{acsim}_{\text{allwi}}$. The 95% confidence intervals are in gray. Each dot represents a pair of repertoires. For each panel, the regression line slope, Mantel test matrix correlation, and Mantel test p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that $\text{acsim}_{\text{allwi}}$ decreases as geographic distance increases. Panels correspond to the: (a) Palindrome; (b) Four-Plus; (c) Slow Increasing; (d) Rapid Increasing; (e) Short; (f) Plus-One; and (g) Regular clans.



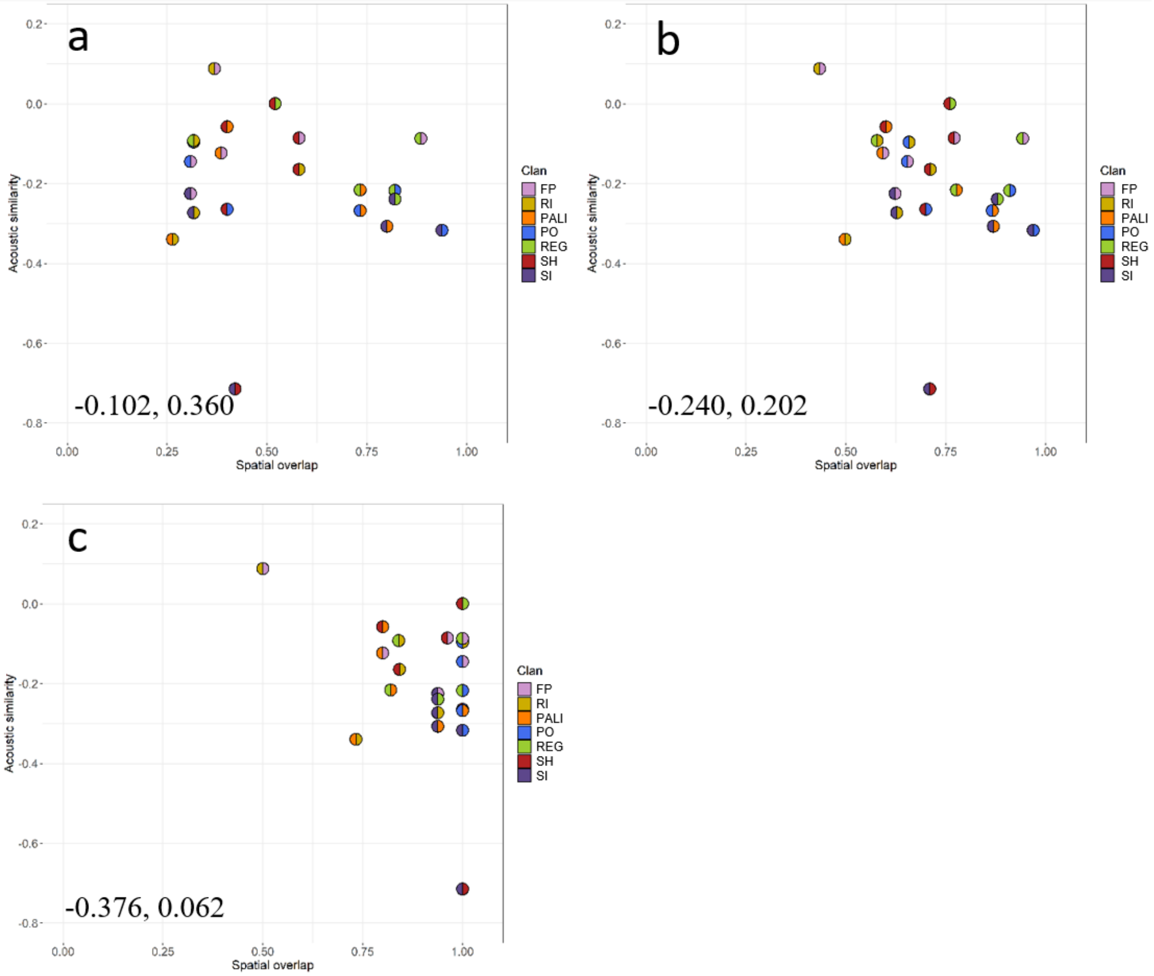
**C4.14:** *Figure S4.10* – Logged geographic distance vs. $\text{acsim}_{\text{diffwi}}$. $\text{acsim}_{\text{diffwi}}$ was calculated as $\text{acsim}_{\text{IDwi}} - \text{acsim}_{\text{nonIDwi}}$. The 95% confidence intervals are in gray. Each dot represents a pair of repertoires. Positive acoustic similarity values indicate that $\text{acsim}_{\text{IDwi}} > \text{acsim}_{\text{nonIDwi}}$, while negative acoustic similarity values indicate that $\text{acsim}_{\text{IDwi}} < \text{acsim}_{\text{nonIDwi}}$. For each panel, the regression line slope, Mantel test matrix correlation, and Mantel test p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that $\text{acsim}_{\text{diffwi}}$ increases as geographic distance increases. Panels correspond to the: (a) Palindrome; (b) Four-Plus; (c) Slow Increasing; (d) Rapid Increasing; (e) Short; (f) Plus-One; and (g) Regular clans.



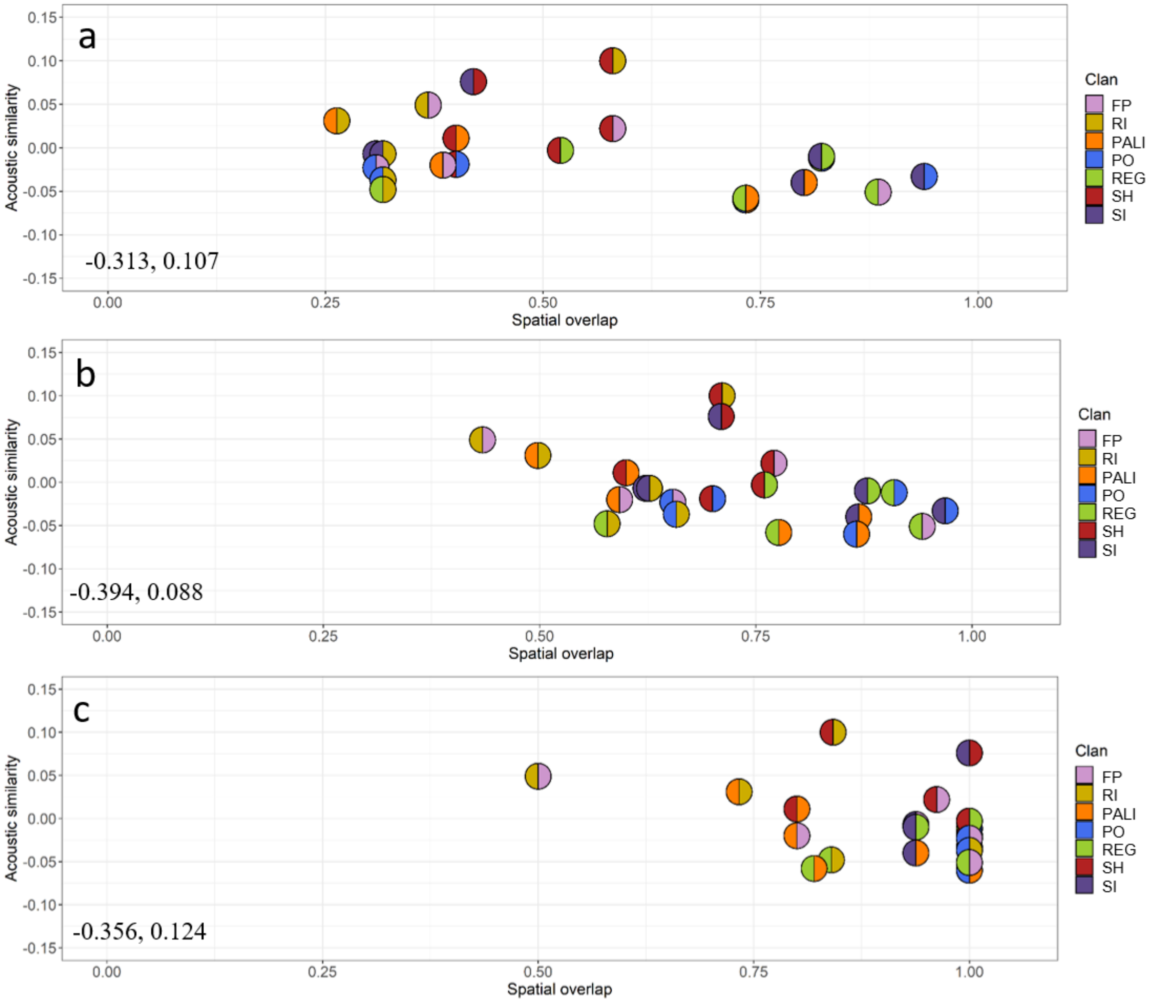
**C4.15:** *Figure S4.11* – Clan spatial overlap vs. $\text{acsim}_{\text{allbt}}$. Each circle represents a pair of clans (see Figure 4.1 for clan abbreviations). For each panel—(a) minimum spatial overlap; (b) mean spatial overlap; and (c) maximum spatial overlap—the Mantel test matrix correlation and p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that $\text{acsim}_{\text{allbt}}$ decreases as clan spatial overlap increases.



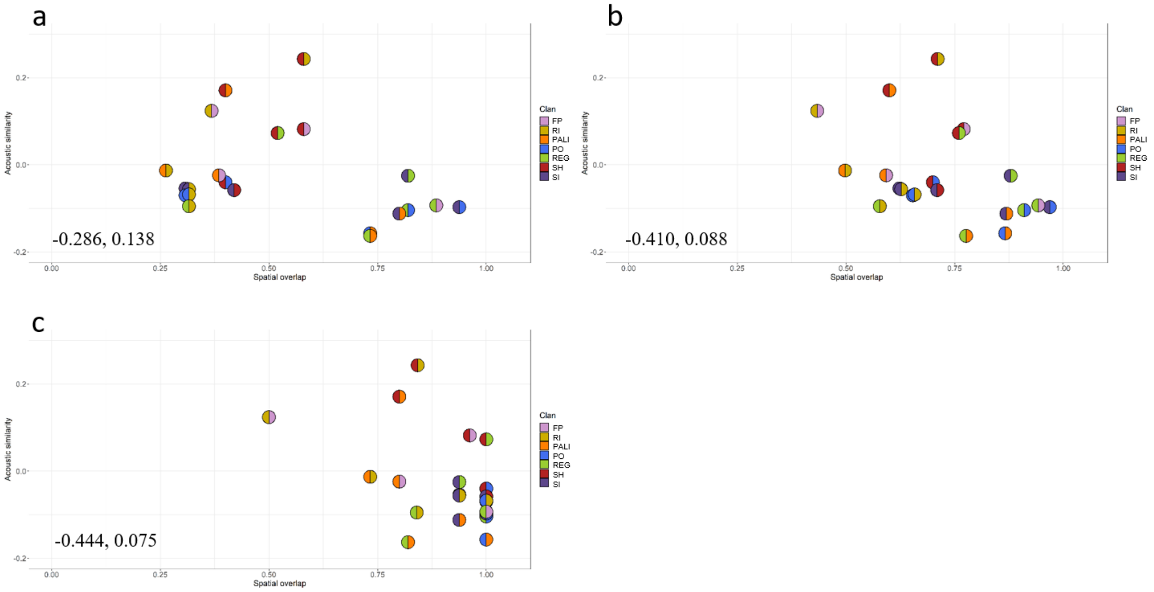
**C4.16:** *Figure S4.12* – Clan spatial overlap vs. acsim$_{\text{diffbt}}$. Each circle represents a pair of clans (see Figure 4.1 for clan abbreviations). For each panel—(a) minimum spatial overlap; (b) mean spatial overlap; and (c) maximum spatial overlap—the Mantel test matrix correlation and p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that acsim$_{\text{diffbt}}$ decreases as clan spatial overlap increases.



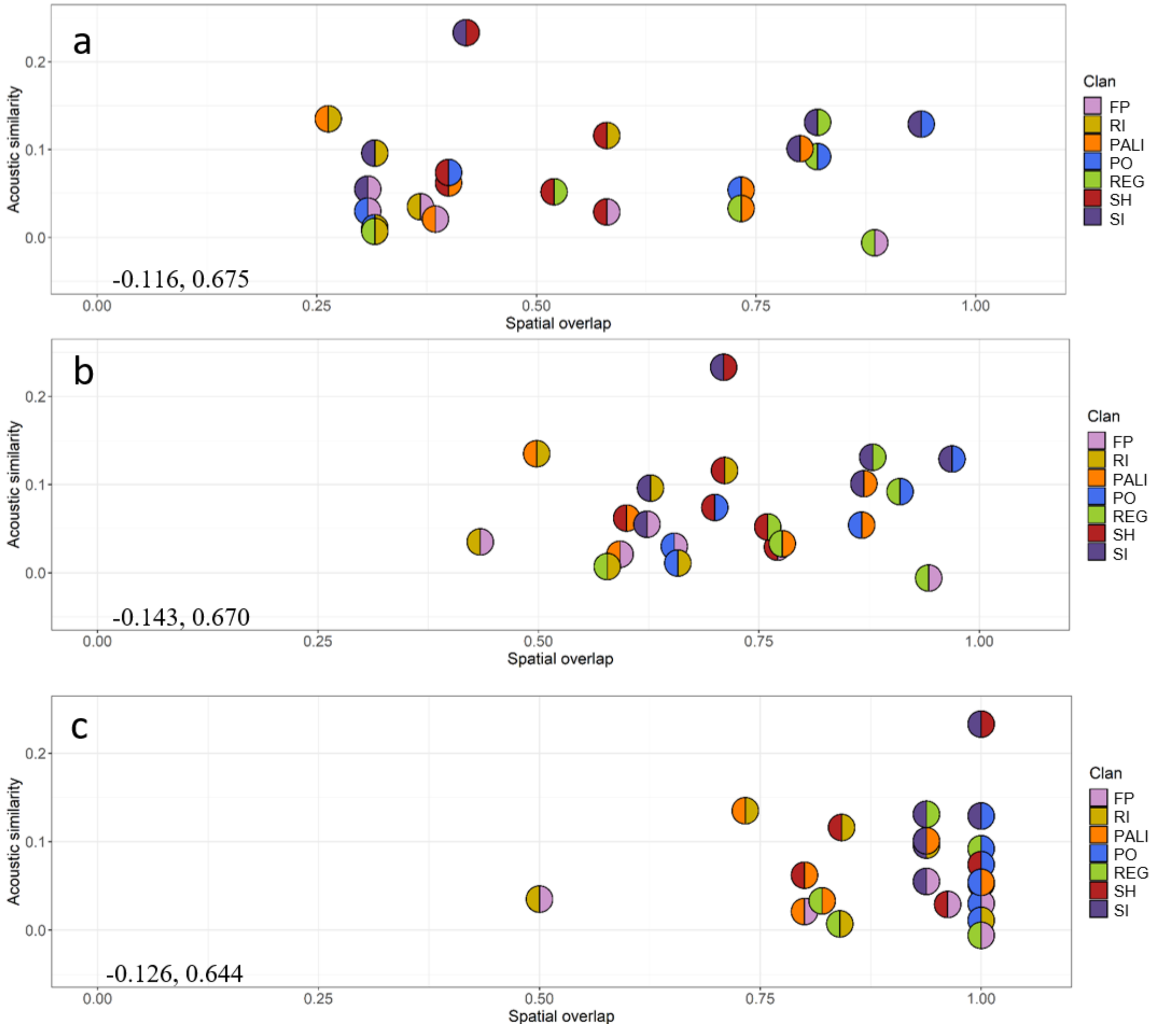
**C4.17:** *Figure S4.13* – Clan spatial overlap vs. $\text{acsim}_{\text{allbtAlt}}$. Each circle represents a pair of clans (see Figure 4.1 for clan abbreviations). For each panel—(a) minimum spatial overlap; (b) mean spatial overlap; and (c) maximum spatial overlap—the Mantel test matrix correlation and p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that $\text{acsim}_{\text{allbtAlt}}$ decreases as clan spatial overlap increases.



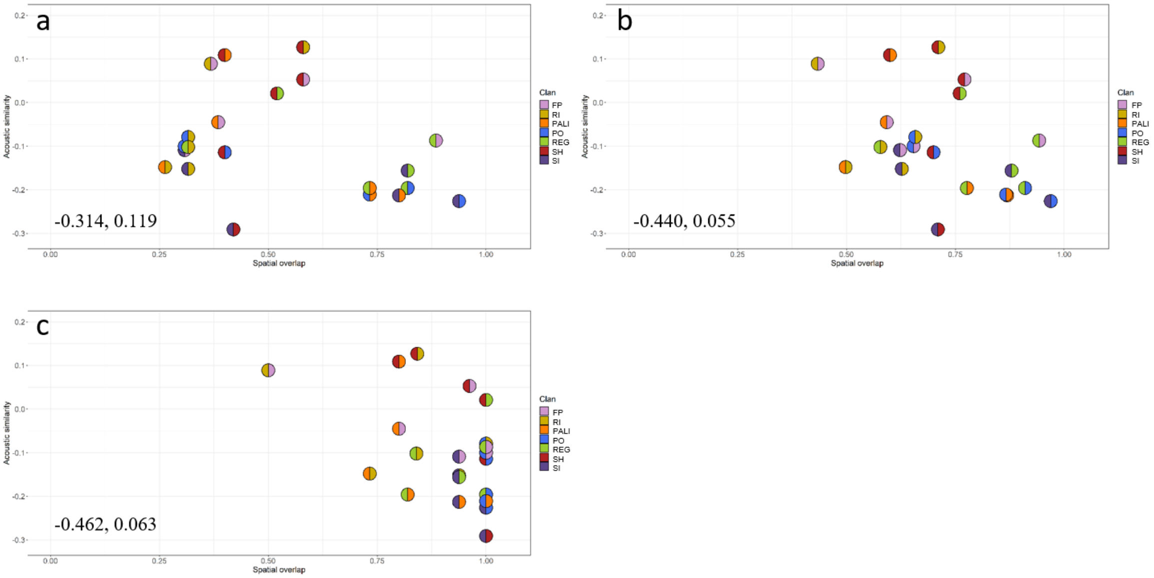
**C4.18:** *Figure S4.14* – Clan spatial overlap vs. $\text{acsim}_{\text{IDbtAlt}}$. Each circle represents a pair of clans (see Figure 4.1 for clan abbreviations). For each panel—(a) minimum spatial overlap; (b) mean spatial overlap; and (c) maximum spatial overlap—the Mantel test matrix correlation and p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that $\text{acsim}_{\text{IDbtAlt}}$ decreases as clan spatial overlap increases.



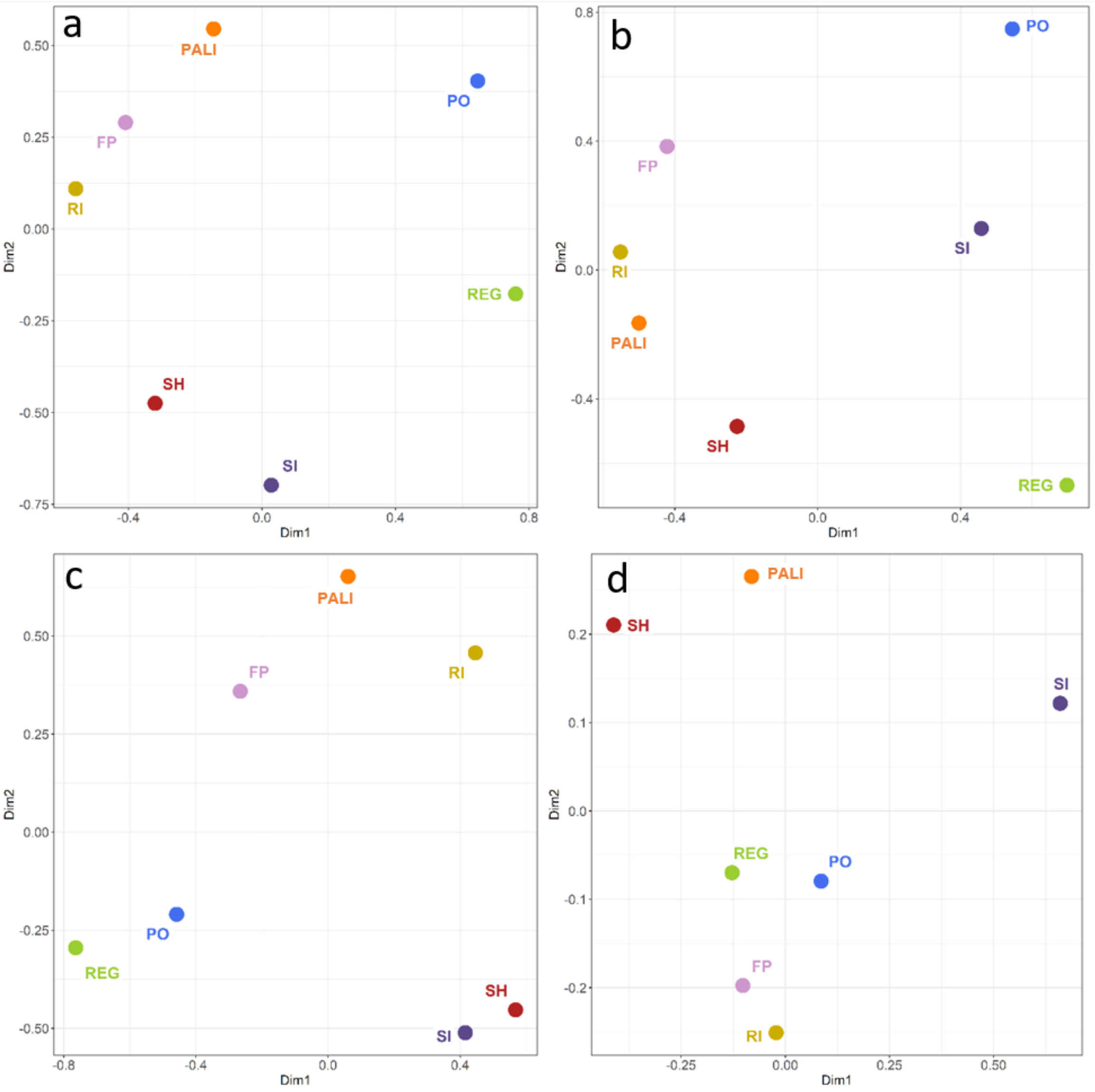
**C4.19:** *Figure S4.15* – Clan spatial overlap vs. acsim$_{\mathrm{nonIDbtAlt}}$. Each circle represents a pair of clans (see Figure 4.1 for clan abbreviations). For each panel—(a) minimum spatial overlap; (b) mean spatial overlap; and (c) maximum spatial overlap—the Mantel test matrix correlation and p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that acsim$_{\mathrm{nonIDbtAlt}}$ decreases as clan spatial overlap increases.



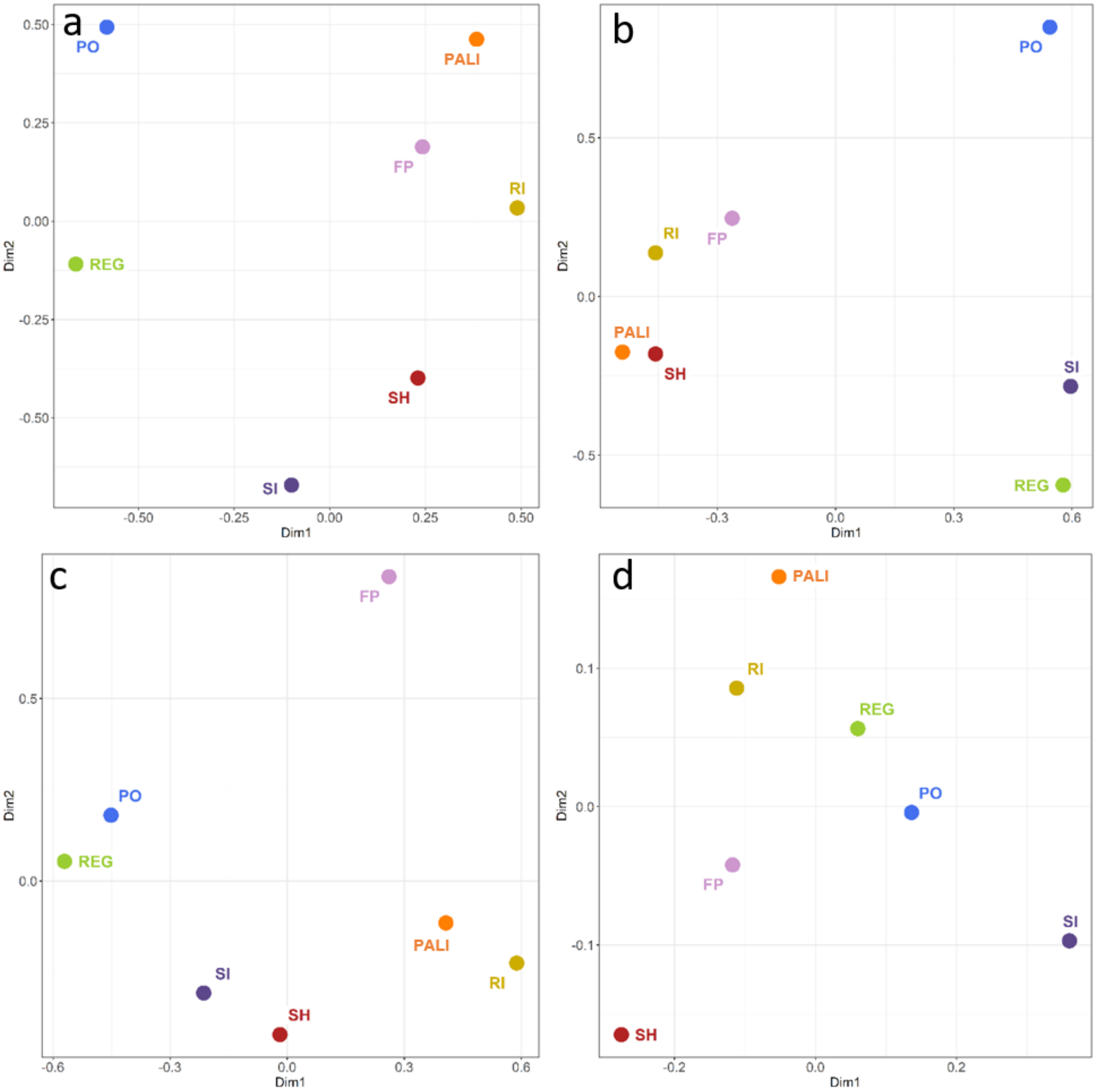
**C4.20:** *Figure S4.16* – Clan spatial overlap vs. acsim$_{\mathrm{diffbtAlt}}$. Each circle represents a pair of clans (see Figure 4.1 for clan abbreviations). For each panel—(a) minimum spatial overlap; (b) mean spatial overlap; and (c) maximum spatial overlap—the Mantel test matrix correlation and p-value are provided in the bottom left corner. Significant p-values are bolded. The Mantel tests are one-tailed against the alternative that acsim$_{\mathrm{diffbtAlt}}$ decreases as clan spatial overlap increases.



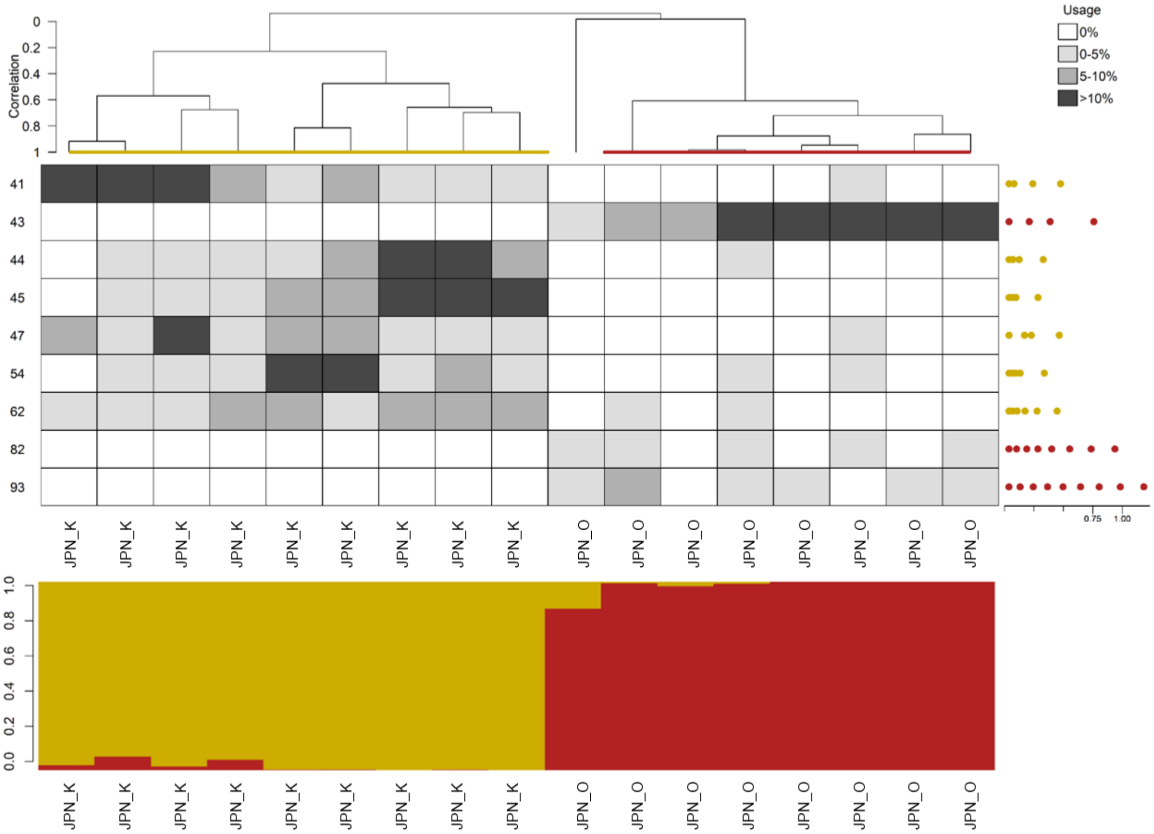
**C4.21:** *Figure S4.17* – Multidimensional scaling plots showing acoustic similarity between clans. Acoustic similarity was calculated as: (a) $\mathrm{acsim}_{\mathrm{allbt}}$ (0.454); (b) $\mathrm{acsim}_{\mathrm{IDbt}}$ (0.445); (c) $\mathrm{acsim}_{\mathrm{nonIDbt}}$ (0.546); and (d) $\mathrm{acsim}_{\mathrm{diffbt}}$ (0.817). Goodness-of-fit values are in parentheses. Note that the axes differ. Plots were created using the ‘cmdscale’ R package.



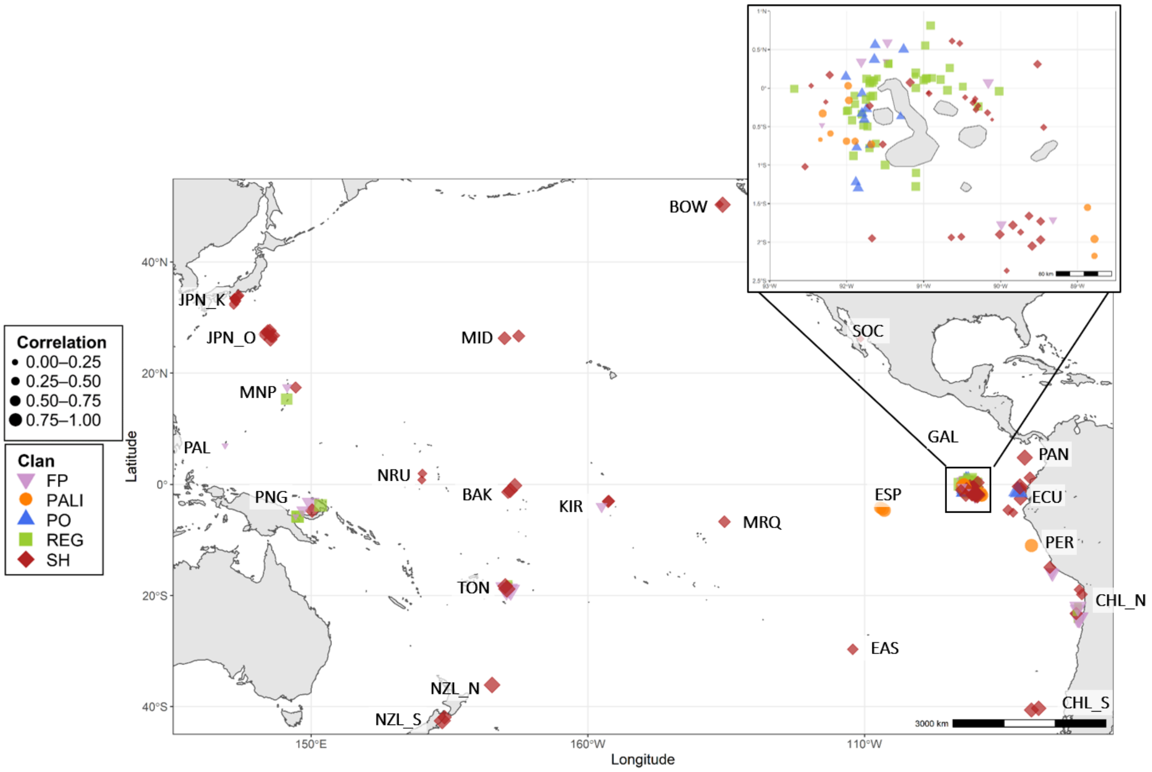
**C4.22:** *Figure S4.18* – Multidimensional scaling plots showing acoustic similarity between clans using alternative method. Acoustic similarity was calculated as: (a) $acsim_{allbtAlt}$ (0.392); (b) $acsim_{IDbtAlt}$ (0.460); (c) $acsim_{nonIDbtAlt}$ (0.421); and (d) $acsim_{diffbtAlt}$ (0.705). Goodness-of-fit values are in parentheses. Note that the axes differ. Plots were created using the 'cmdscale' R package.



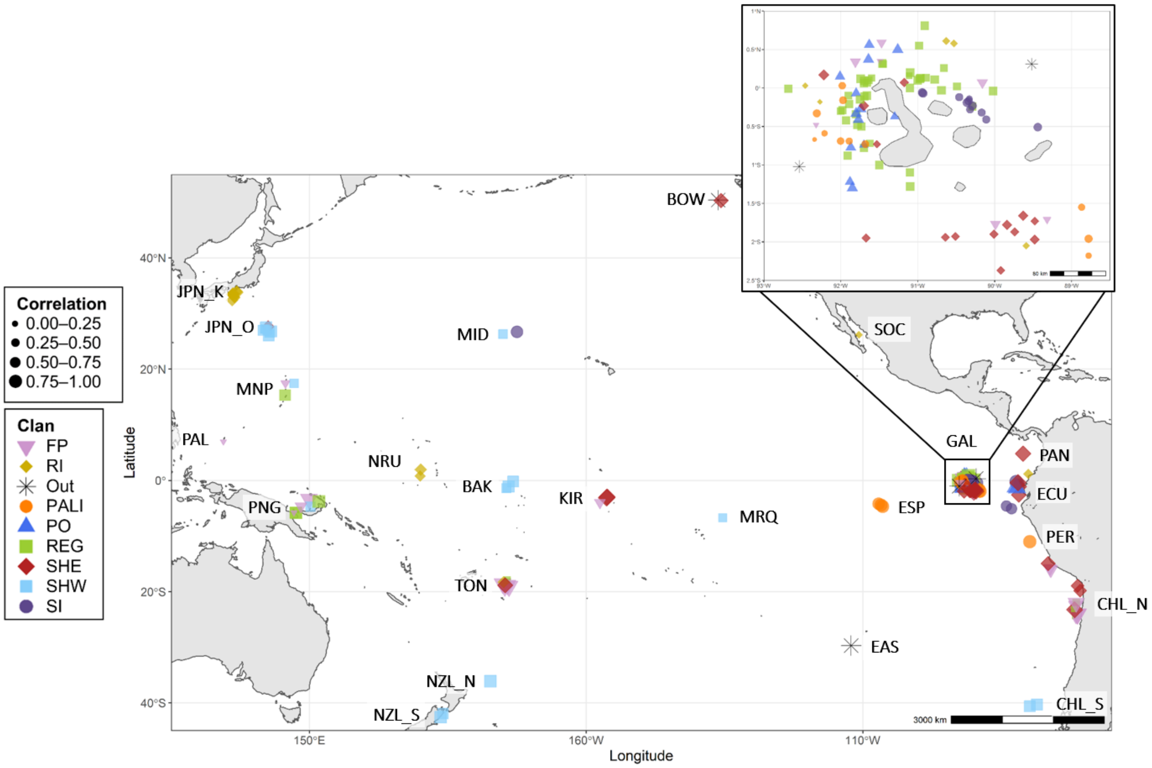
**C4.23:** *Figure S4.19* – Example tree with identity coda types showing the division of Kumano coast and Ogasawara Islands coda repertoires. This tree was constructed using *critfact*=15 and *minrep*=5 and the division of repertoires into clans was replicated at very high levels of *critfact* (e.g. *critfact*=310). The trends in coda usage seen here mirror those reported in (Amano et al., 2014), with Kumano coast (JPN_K) whales (gold) making identity codas with shorter durations than Ogasawara Islands (JPN_O) whales (red; see Figure 4 in Amano et al. 2014). See Figure 4.2 for additional details.



**C4.24:** *Figure S4.20* – Composite map of sperm whale clan distribution across 23 regions in the Pacific Ocean using the 5-clan tree clan designations. See Table S4.1 for region abbreviations and Figure 4.1 for additional details.



**C4.25:** *Figure S4.21* – Composite map of sperm whale clan distribution across 23 regions in the Pacific Ocean using the 8-clan tree clan designations. See Table S4.1 for region abbreviations and Figure 4.1 for additional details. Additional abbreviations are: SHE=Short East clan, SHW=Short West clan, and Out=Outlier (i.e. a repertoire that was not assigned to a clan).



**C4.26:** *Table S4.1* – Pacific region abbreviations and recording metadata. Asterisks denote regions for which codas had previously been extracted and described, although additional codas from some of those regions were extracted in the present study as well. Citations with information on the recording equipment and field methods are provided in the 'Recording details' column. If citations were unavailable, coauthors provided the relevant details (see 'Coauthor(s)' column for corresponding initials). Codas from some regions, such as Bowie Seamount, were recorded using stationary, bottom mounted hydrophones which enabled precise localization for each repertoire. In other regions (e.g. Kumano coast of Japan, Ogasawara Islands of Japan), recording locations were given as ranges of latitude and longitude (spanning 0.1 to 0.5 decimal degrees). No GPS was available for a small number of southern New Zealand codas recorded in 1992 (n=36) off Kaikōura, so they were assigned the average GPS of codas recorded off Kaikōura in 2009 and 2010 by MF. Different recording systems were used across regions and years, but this does not affect our measurements of the temporal patterns of codas. See text for additional details.

| **Region** | **Recording year(s)** | **Recording details** | **Coauthor(s)** |
|---|---|---|---|
| Baker Island (BAK)* | 1992 | (Weilgart & Whitehead, 1997) | LR, LW, HW |
| Bowie Seamount (BOW) | 2011, 2013 | (Riera et al., 2016) | TD-V, JP |
| Northern Chile (CHL_N)* | 1993, 2000 | (Rendell & Whitehead, 2003b; Weilgart & Whitehead, 1997) | LR, LW, HW |
| Southern Chile (CHL_S)* | 1993 | (Weilgart & Whitehead, 1997) | LW, HW |
| Easter Island (EAS)* | 1993 | (Rendell & Whitehead, 2003b; Weilgart & Whitehead, 1997) | LR, LW, HW |
| Ecuador (ECU)* | 1985, 1991, 1993 | (Rendell & Whitehead, 2003b; Weilgart & Whitehead, 1997) | TAH, LR, LW, HW |
| Equatorial South Pacific (ESP) | 2000 | (Godard et al., 2003) | CJ, IK, RP, AR |
| Galápagos Islands (GAL)* | 1978, 1985, 1987, 1989, 1991, 1995, 1999, 2000, 2013, 2014 | (Cantor et al., 2016; Godard et al., 2003; Rendell & Whitehead, 2003b; Sayigh et al., 2016; Weilgart & Whitehead, 1997) | MC, TAH, IK, LR, AR, LW, HW |

*Continued on next page*

| **Region** | **Recording year(s)** | **Recording details** | **Coauthor(s)** |
|---|---|---|---|
| Kumano coast of Japan (JPN_K)* | 2004, 2007, 2008 | (Amano et al., 2014) | MA |
| Ogasawara Islands of Japan (JPN_O)* | 1994, 1995, 1996, 2006, 2007, 2008 | (Amano et al., 2014) | MA |
| Kiribati (KIR)* | 1992 | (Weilgart & Whitehead, 1997) | LW, HW |
| Midway Atoll (MID) | 2013, 2017 | (Barkley et al., 2019) | YB, KM, EO |
| Mariana Islands (MNP) | 2007 | (Norris et al., 2012; Waller et al., 2007) | OA, EF, CAH-W, TN |
| Marquesas Islands (MRQ) | 2000 | (Godard et al., 2003) | CJ, IK, RP, AR |
| Nauru (NRU) | 2001 | (Godard et al., 2003) | CJ, IK, RP, AR |
| Northern New Zealand (NZL_N)* | 1993 | (Weilgart & Whitehead, 1997) | LW, HW |
| Southern New Zealand (NZL_S) | 1992, 1999, 2000, 2006, 2007, 2009, 2010, 2013, 2015 | (Fernandes, 2016; Gordon et al., 1992; Guerra et al., 2020; Jaquet et al., 2001; Miller et al., 2013; Rhinelander & Dawson, 2004) | MF, JG, MG, LH, ES |

*Continued on next page*

| Region | Recording year(s) | Recording details | Coauthor(s) |
|---|---|---|---|
| Palau (PAL) | 2012 | Recordings were made from a 50 m survey vessel (*Endless Summer*) using a towed hydrophone array. The array consisted of a 70 m cable with two hydrophone elements (Reson TC4013) spaced 3 m apart. Both array channels were passed through a tunable high pass filter (Ecologic HP/27 ST Magrec Stereo Monitor Box) to reduce water and vessel noise. Corner frequencies of the high-pass filter were set between 500 Hz and 1 kHz depending on noise, and the gain was set between 10 to 20 decibels. The signal was split from the high-pass filter and fed into a PC digital interface (MOTU Traveler) and multi-track digital recorder (Tascam DR-680). Files were recorded at a 192 kHz sampling rate, 24-bit resolution. | OA, EF, CAH-W, TN |
| Panama (PAN)* | 1992 | (Weilgart & Whitehead, 1997) | LW, HW |
| Peru (PER)* | 1993 | (Pavan et al., 2000) | TAH, LW, HW |
| Papua New Guinea (PNG) | 2001 | (Godard et al., 2003) | CJ, IK, RP, AR |
| Sea of Cortez (SOC) | 1999 | (Godard et al., 2003) | CJ, IK, RP, AR |
| Tonga (TON) | 1992, 2003 | (Weilgart & Whitehead, 1997)<br>Recordings in 2003 were made from a 15.3 m sailing vessel (*RV Catalyst*) using a towed hydrophone array. The array consisted of a 150 m cable with three hydrophone elements. Sperm whale clicks were detected in recordings using ‘Rainbow Click’ (Gillespie, 1997). Codas were then identified and measured manually. | RA, ES, LW, HW |

**C4.27:** *Table S4.2* – Extracted coda and repertoire information for each Pacific region. A repertoire is comprised of all the codas with 3–10 clicks recorded in a specific region on a single day. See Table S4.1 for region abbreviations. All 3–10-click codas (third column) were included in the call classification stage of IDcall, but only repertoires with at least 25 codas (fourth column) were included in the hierarchical clustering stage. The 'Year(s)' and 'Number of repertoires' columns are based only on repertoires with at least 25 codas, which is why some years present in Table S4.1 are absent here.

| **Region** | **Total number of codas (2–29 clicks)** | **Number of 3–10 click codas (% of total)** | **Number of 3–10 click codas from repertoires with ≥25 codas (% of total)** | **Year** | **Number of repertoires** |
|---|---|---|---|---|---|
| BAK | 278 | 272 (97.8%) | 272 (97.8%) | 1992 | 3 |
| BOW | 133 | 126 (94.7%) | 126 (94.7%) | 2011 | 1 |
| | | | | 2013 | 1 |
| CHL_N | 5,538 | 5,523 (99.7%) | 5,523 (99.7%) | 1993 | 1 |
| | | | | 2000 | 11 |
| CHL_S | 183 | 183 (100%) | 183 (100%) | 1993 | 2 |
| EAS | 90 | 90 (100%) | 90 (100%) | 1993 | 1 |
| ECU | 779 | 774 (99.4%) | 774 (99.4%) | 1985 | 1 |
| | | | | 1991 | 5 |
| | | | | 1993 | 3 |
| ESP | 263 | 237 (90%) | 215 (81.7%) | 2000 | 3 |
| GAL | 10,959 | 10,507 (95.9%) | 10,232 (93.4%) | 1978 | 1 |
| | | | | 1985 | 19 |
| | | | | 1987 | 24 |
| | | | | 1989 | 11 |
| | | | | 1991 | 3 |
| | | | | 1995 | 8 |
| | | | | 1999 | 10 |
| | | | | 2000 | 1 |
| | | | | 2013 | 13 |
| | | | | 2014 | 14 |
| JPN_K | 836 | 829 (99.2%) | 768 (91.9%) | 2007 | 6 |
| | | | | 2008 | 3 |
| JPN_O | 985 | 846 (85.9%) | 697 (70.8%) | 2006 | 1 |
| | | | | 2007 | 2 |
| | | | | 2008 | 5 |
| KIR | 511 | 511 (100%) | 511 (100%) | 1992 | 3 |
| MID | 250 | 239 (95.6%) | 239 (95.6%) | 2013 | 1 |
| | | | | 2017 | 1 |
| MNP | 278 | 164 (59.0%) | 149 (53.6%) | 2007 | 3 |
| MRQ | 61 | 60 (98.4%) | 60 (98.4%) | 2000 | 1 |

*Continued on next page*

| **Region** | **Total number of codas (2–29 clicks)** | **Number of 3–10 click codas (% of total)** | **Number of 3–10 clicks from repertoires with ≥ 25 codas (% of total)** | **Year** | **Number of repertoires** |
|---|---|---|---|---|---|
| NRU | 89 | 88 (98.9%) | 88 (98.9%) | 2001 | 2 |
| NZL_N | 44 | 44 (100%) | 44 (100%) | 1993 | 1 |
| NZL_S | 157 | 156 (99.4%) | 96 (61.1%) | 1992<br>2010 | 1<br>1 |
| PAL | 34 | 31 (91.2%) | 31 (91.2%) | 2012 | 1 |
| PAN | 191 | 191 (100%) | 191 (100%) | 1992 | 1 |
| PER | 672 | 665 (99.0%) | 665 (99.0%) | 1993 | 6 |
| PNG | 450 | 443 (98.4%) | 443 (98.4%) | 2001 | 7 |
| SOC | 207 | 202 (97.6%) | 201 (97.1%) | 1999 | 1 |
| TON | 1,249 | 1,248 (99.9%) | 1,232 (98.6%) | 1992<br>2003 | 1<br>7 |
| **Totals** | **24,237** | **23,429 (96.7%)** | **22,829 (94.2%)** | **21** | **191** |

**C4.28:** *Table S4.3* – Dendrogram features as *critfact* varied with *minrep* kept constant. *critfact* varied from 3 to 20 and *minrep* was 15. See Figures 4.1/S4.21 for clan abbreviations

| ***critfact*** | **Number of clans** | **Clans** | **Total number of identity codas** | **Number of repertoires not assigned to a clan** |
|---|---|---|---|---|
| 3 | 8 | PALI, FP, SI, RI, SHW, SHE, PO, REG | 34 | 4 |
| 4 | | | 35 | |
| 5 | 7 | PALI, FP, SI, RI, SH, PO, REG | 32 | 0 |
| 6 | 8 | PALI, FP, SI, RI, SHW, SHE, PO, REG | 33 | 3 |
| 7 | 4 | PALI, SI, PO, REG | 25 | 95 |
| 8 | | | 24 | |
| 9 | | | 23 | |
| 10 | | | | |
| 11 | | | | |
| 12 | 5 | PALI, FP, SH, PO, REG | 23 | 0 |
| 13 | | | 24 | |
| 14 | | | 23 | |
| 15 | | | | |
| 16 | | | 21 | |
| 17 | | | | |
| 18 | | | | |
| 19 | | | | |
| 20 | | | 20 | |

**C4.29:** *Table S4.4* – Summary of clans and identity codas in an exemplar 5-clan tree. Tree was created using *critfact*=12 and *minrep*=15. See Figure S4.20 for clan abbreviations. This Table is like Table 4.1 except the 'Large Short' clan encompasses the SH, RI, and SI clans from Table 4.1 and some identity codas have changed. See Table 4.1 for additional information.

| **Clan** | **Number of repertoires** | **Within-clan correlation (mean ± SD)** | **Number of identity codas** | **Identity coda types** | | |
|---|---|---|---|---|---|---|
| | | | | *Number of clicks* | *Numeric code* | *Type names* |
| PALI | 15 | 0.699 ± 0.197 | 7 | 4<br>7<br>9<br>10 | 411, 412<br>73, 711<br>94, 95<br>102 | 2+2, 2+2<br>3+1+3, 3+1+3<br>9I, 9I<br>10I |
| FP | 26 | 0.445 ± 0.314 | 2 | 6 | 611, 614 | 4+1++1, 4+1++1 |
| Large Short | 85 | 0.367 ± 0.210 | 1 | 3 | 33 | 1+2 |
| PO | 15 | 0.854 ± 0.123 | 5 | 5<br>6<br>7<br>8 | 515<br>68, 613<br>712<br>89 | 4+1<br>5+1, 5+1<br>7R<br>8R |
| REG | 50 | 0.763 ± 0.243 | 8 | 5<br>6<br>7<br>8<br>9<br>10 | 512<br>62<br>71, 74<br>81<br>92, 96<br>108 | 5R<br>6R<br>7R, 7R<br>8R<br>9R, 9R<br>10R |

**C4.30:** *Table S4.5* – Summary of clans and identity codas in an exemplar 8-clan tree. Tree was created using *critfact*=6 and *minrep*=15. See Figure S4.21 for clan abbreviations. This Table is identical to Table 4.1 except the Table 4.1 SH clan is divided into SHW and SHE here. Additionally, four repertoires that were assigned to the SH clan in the 7-clan tree were not assigned to a clan in the 8-clan tree. See Table 4.1 for additional information.

| **Clan** | **Number of repertoires** | **Within-clan correlation (mean ± SD)** | **Number of identity codas** | **Identity coda types** | | |
|---|---|---|---|---|---|---|
| | | | | *Number of clicks* | *Numeric code* | *Type names* |
| PALI | 15 | 0.699 ± 0.197 | 9 | 4 | 48, 411, 412 | 1+1++2, 2+2, 2+2 |
| | | | | 7 | 73, 711 | 3+1+3, 3+1+3 |
| | | | | 8 | 82 | 8I |
| | | | | 9 | 94, 95 | 9I, 9I |
| | | | | 10 | 102 | 10I |
| FP | 26 | 0.445 ± 0.314 | 2 | 6 | 611, 614 | 4+1++1, 4+1++1 |
| SI | 16 | 0.730 ± 0.175 | 3 | 3 | 39 | 2+1 |
| | | | | 4 | 413 | 4I |
| | | | | 7 | 69 | 6I |
| RI | 19 | 0.652 ± 0.185 | 2 | 4 | 45 | 4I |
| | | | | 7 | 715 | 7I |
| SHW | 19 | 0.608 ± 0.187 | 1 | 3 | 310 | 3R |
| SHE | 27 | 0.625 ± 0.166 | 1 | 3 | 313 | 3R |
| PO | 15 | 0.854 ± 0.123 | 6 | 5 | 515 | 4+1 |
| | | | | 6 | 68, 613 | 5+1, 5+1 |
| | | | | 7 | 79, 712 | 6+1, 7R |
| | | | | 8 | 89 | 8R |
| REG | 50 | 0.763 ± 0.243 | 9 | 5 | 51, 512 | 5R, 5R |
| | | | | 6 | 62 | 6R |
| | | | | 7 | 71, 74 | 7R, 7R |
| | | | | 8 | 81 | 8R |
| | | | | 9 | 92, 96 | 9R, 9R |
| | | | | 10 | 108 | 10R |

**C4.31:** *Table S4.6* – Number of detected clans per region with regional sampling effort. ‘Number of repertoires’ refers to repertoires with at least 25 codas and ‘Number of codas’ refers to codas with 3–10 clicks in those repertoires. Regions are ordered from most to least clans, then from most to least repertoires, and finally from most to least codas. See Table S4.1 for region abbreviations.

| **Region** | **Number of clans** | **Number of repertoires** | **Number of codas** |
|---|---|---|---|
| GAL | 7 | 104 | 10,232 |
| ECU | 5 | 9 | 774 |
| TON | 4 | 8 | 1,232 |
| PER | 4 | 6 | 665 |
| CHL_N | 3 | 12 | 5,523 |
| PNG | 3 | 7 | 443 |
| MNP | 3 | 3 | 149 |
| KIR | 2 | 3 | 511 |
| MID | 2 | 2 | 239 |
| JPN_K | 1 | 9 | 768 |
| JPN_O | 1 | 8 | 697 |
| BAK | 1 | 3 | 272 |
| ESP | 1 | 3 | 215 |
| CHL_S | 1 | 2 | 183 |
| BOW | 1 | 2 | 126 |
| NZL_S | 1 | 2 | 96 |
| NRU | 1 | 2 | 88 |
| EAS | 1 | 1 | 90 |
| SOC | 1 | 1 | 201 |
| PAN | 1 | 1 | 191 |
| MRQ | 1 | 1 | 60 |
| NZL_N | 1 | 1 | 44 |
| PAL | 1 | 1 | 31 |

**C4.32:** *Table S4.7* – Clan spatial overlap values for the seven Pacific clans. The matrix gives the proportion of the row clan’s repertoires that were recorded within 1,000 km of at least one of the column clan’s repertoires. Note that the matrix is asymmetric, as a clan found in only one region might overlap completely with a clan that spans the ocean, while the inverse is not true (e.g. see PO/SH clans). See Figure 4.1 for clan abbreviations.

| **Clan** | **FP** | **RI** | **PALI** | **PO** | **REG** | **SH** | **SI** |
|---|---|---|---|---|---|---|---|
| **FP** | 1 | 0.500 | 0.385 | 0.308 | 0.885 | 0.962 | 0.308 |
| **RI** | 0.368 | 1 | 0.263 | 0.316 | 0.316 | 0.842 | 0.316 |
| **PALI** | 0.800 | 0.733 | 1 | 0.733 | 0.733 | 0.800 | 0.800 |
| **PO** | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **REG** | 1 | 0.840 | 0.820 | 0.820 | 1 | 1 | 0.820 |
| **SH** | 0.580 | 0.580 | 0.400 | 0.400 | 0.520 | 1 | 0.420 |
| **SI** | 0.938 | 0.938 | 0.938 | 0.938 | 0.938 | 1 | 1 |

**C4.33:** *Table S4.8* – Convention on Migratory Species Range States in which sperm whales from different Pacific clans have been identified. This is an updated version of a table from the Convention on Migratory Species (2017). Data are from X: Rendell & Whitehead (2003b), Y: Cantor et al. (2016), or Z: present study (see Figure S4.8). Our results generally mirrored and extended those of Cantor et al. (2016) and Rendell & Whitehead (2003b), but we did not find evidence of the Plus-One clan off Panama or the Regular clan off Peru, as Rendell & Whitehead (2003b) did. For clan abbreviations, see Figure 4.1.

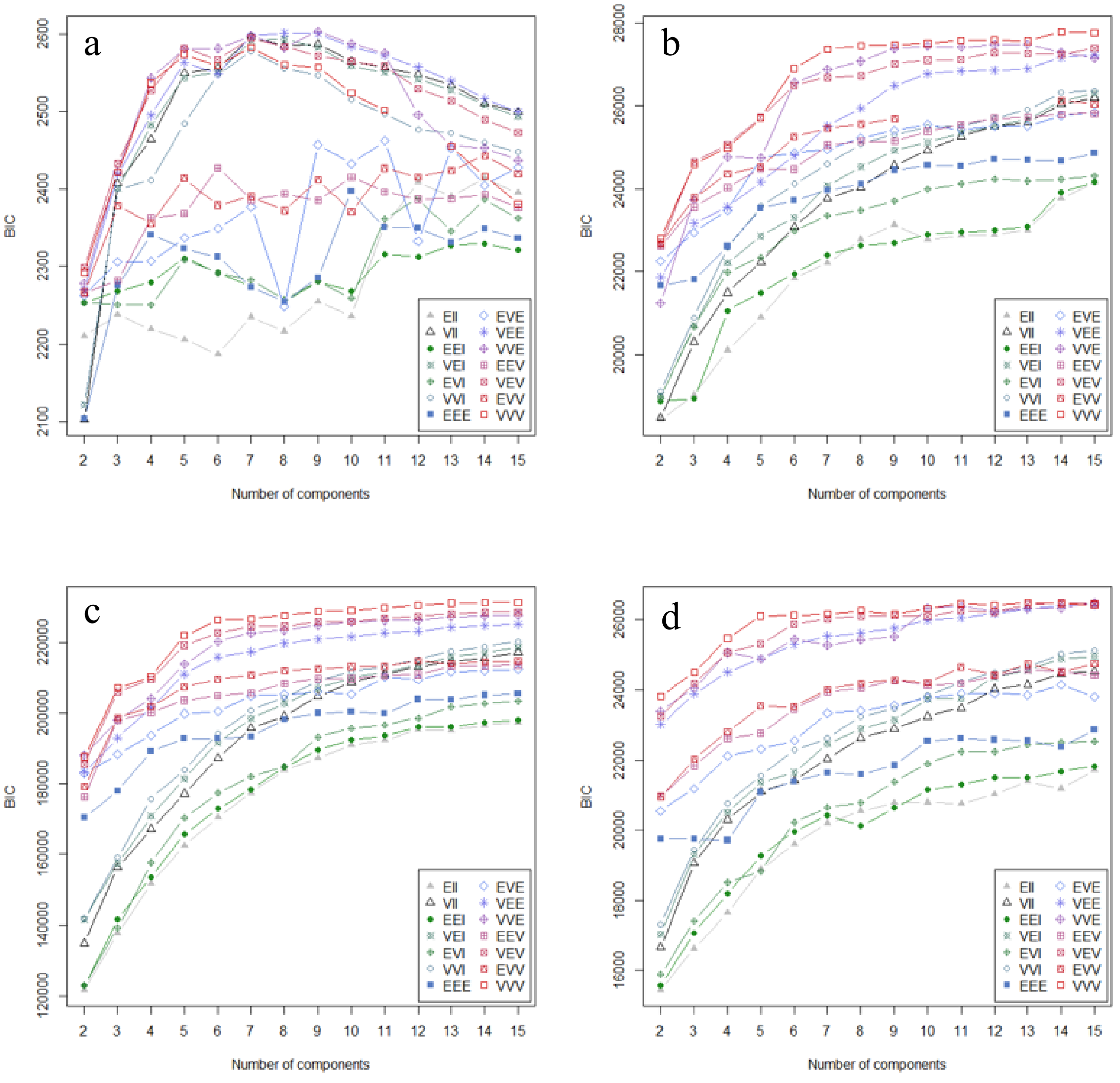
| | **Panama** | **Ecuador**[43] | **Peru** | **Chile** |
|---|---|---|---|---|
| **REG** | | X, Z | X | X, Z |
| **PO** | X | X, Z | | |
| **SH** | X, Z | X, Y, Z | X, Z | X, Z |
| **FP** | | Y, Z | X, Z | X, Z |
| **PALI** | | Z | Z | |
| **RI** | | Z | | |
| **SI** | | Z | Z | |

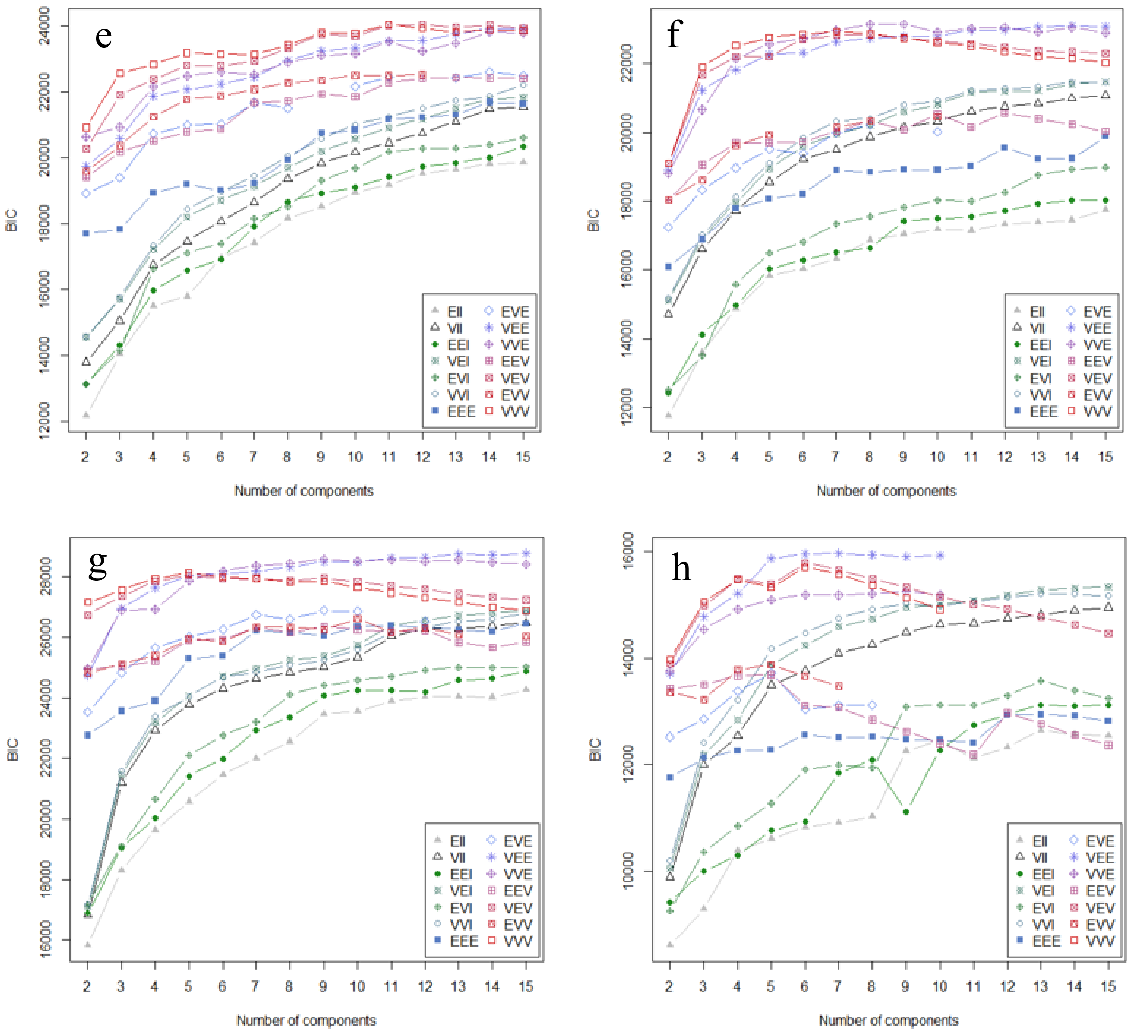
---

[43] Includes mainland Ecuador and the Galápagos Islands.

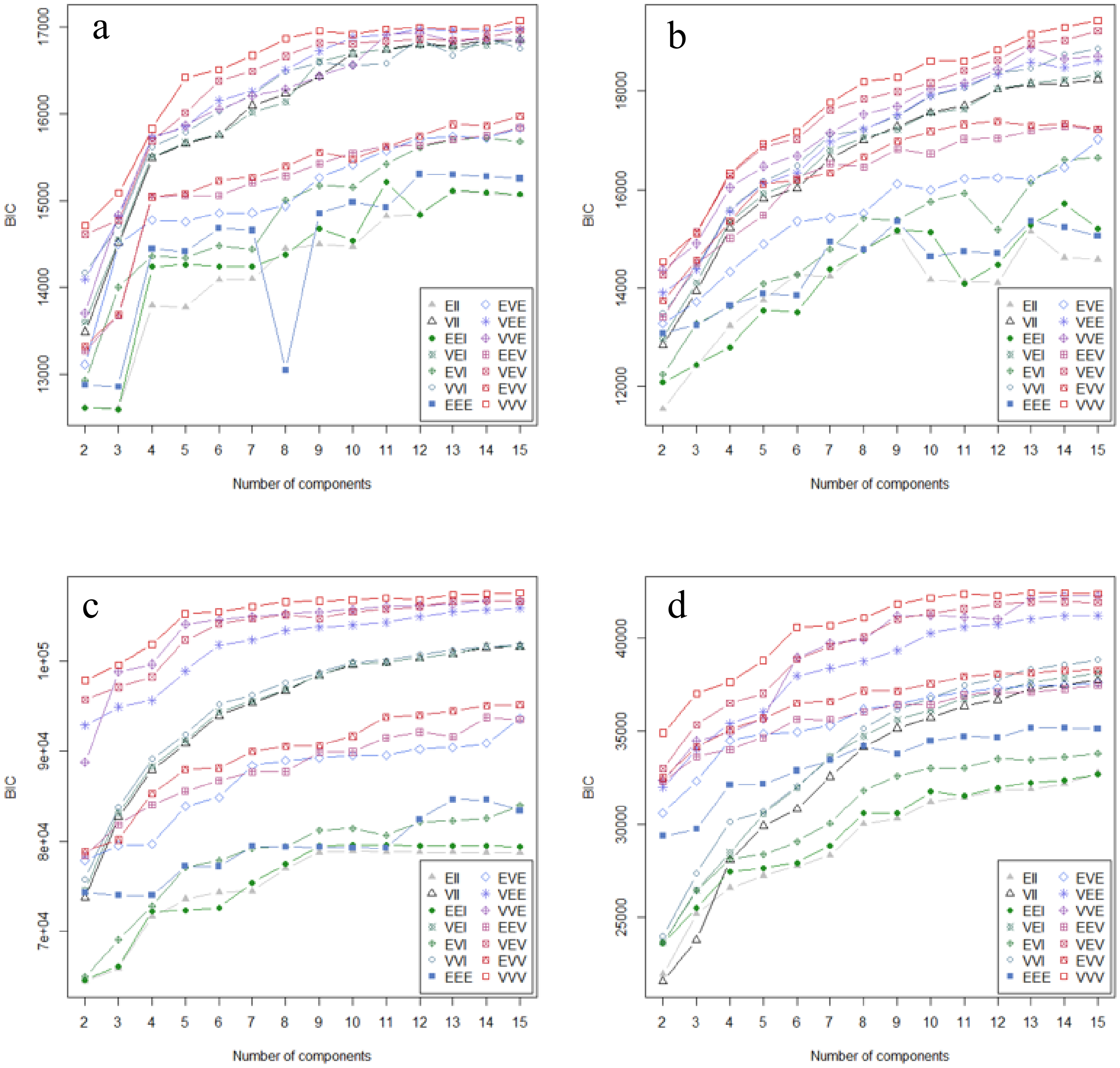
# APPENDIX D – CHAPTER 5 SUPPLEMENT

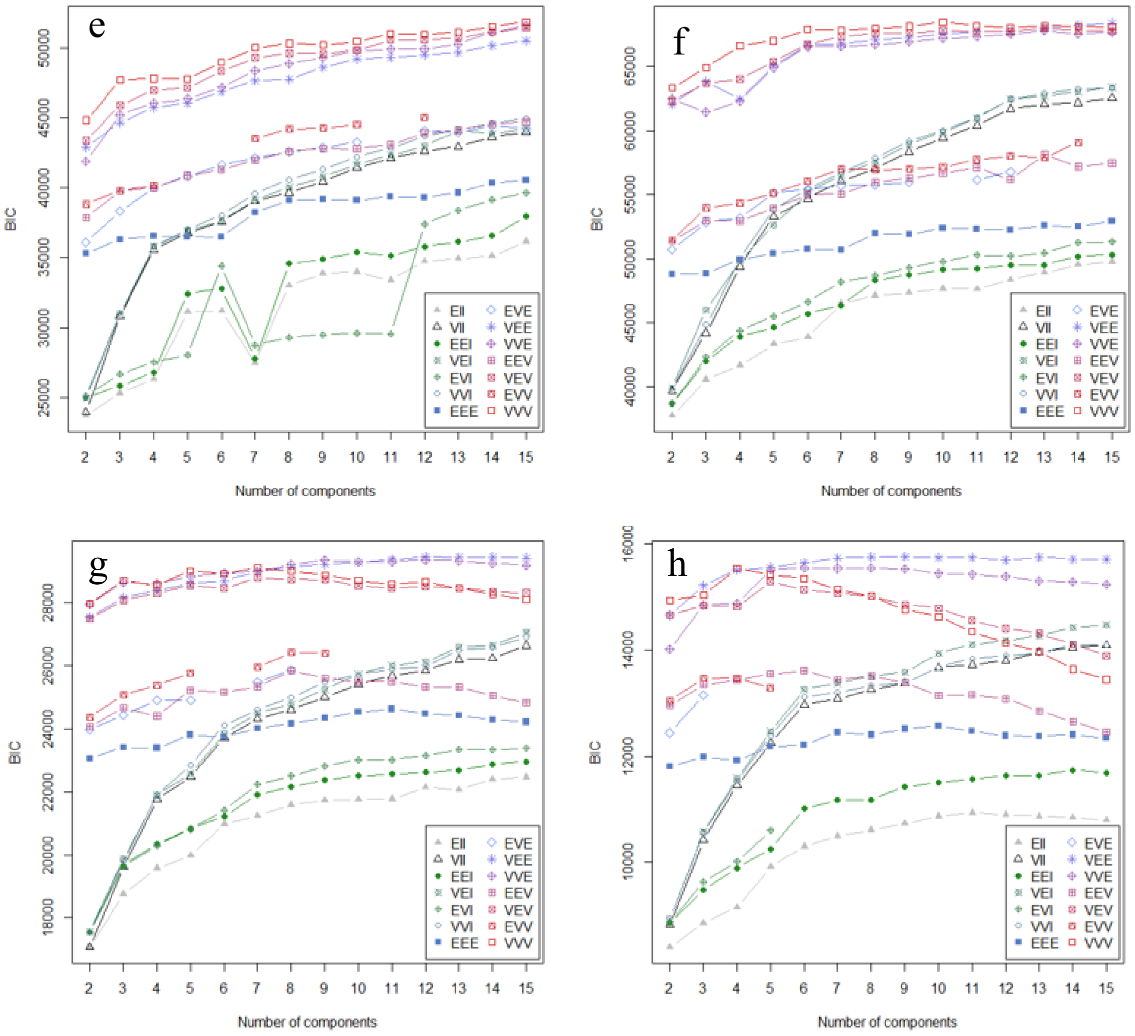
**D5.1:** *Figure S5.1* – Results of preliminary Atlantic/Mediterranean data exploration using mclust. Plots show the BIC value for each of 14 models (Punzo & McNicholas, 2016) when fitting 2–15 mixture components to 3–10-click codas (panels a–h). A higher BIC indicates a better fit.



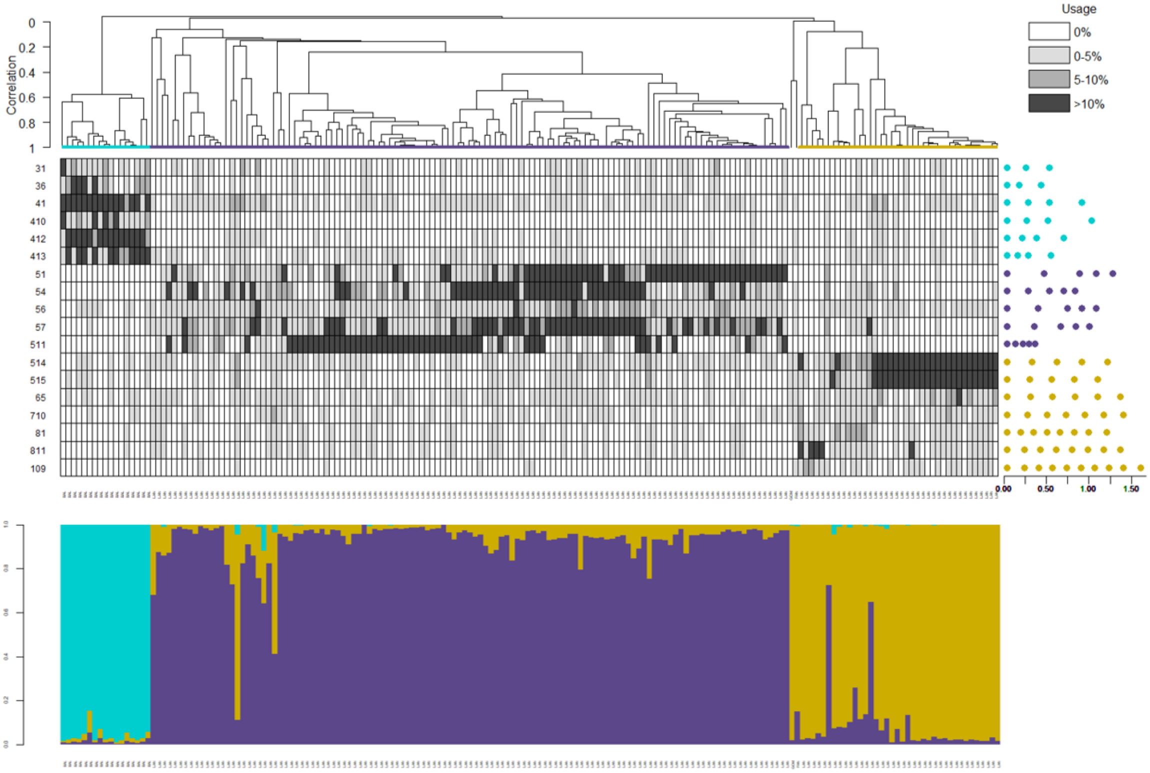

e
f
g
h
BIC
Number of components
2 3 4 5 6 7 8 9 10 11 12 13 14 15
12000 14000 16000 18000 20000 22000 24000
12000 14000 16000 18000 20000 22000
16000 18000 20000 22000 24000 26000 28000
10000 12000 14000 16000
EII
VII
EEI
VEI
EVI
VVI
EEE
EVE
VEE
VVE
EEV
VEV
EVV
VVV

**D5.2:** *Figure S5.2* – Results of preliminary Pacific data exploration using mclust. Plots show the BIC value for each of 14 models (Punzo & McNicholas, 2016) when fitting 2–15 mixture components to 3–10-click codas (panels a–h). A higher BIC indicates a better fit.

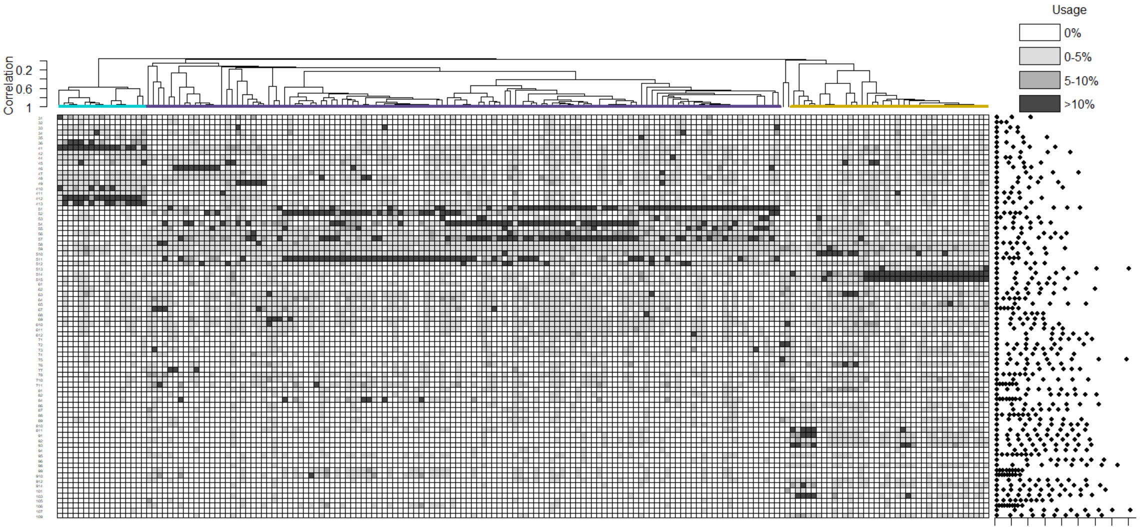



e
f
g
h
BIC
Number of components
2 3 4 5 6 7 8 9 10 11 12 13 14 15
25000 30000 35000 40000 45000 50000
40000 45000 50000 55000 60000 65000
18000 20000 22000 24000 26000 28000
10000 12000 14000 16000
EII
VII
EEI
VEI
EVI
VVI
EEE
EVE
VEE
VVE
EEV
VEV
EVV
VVV

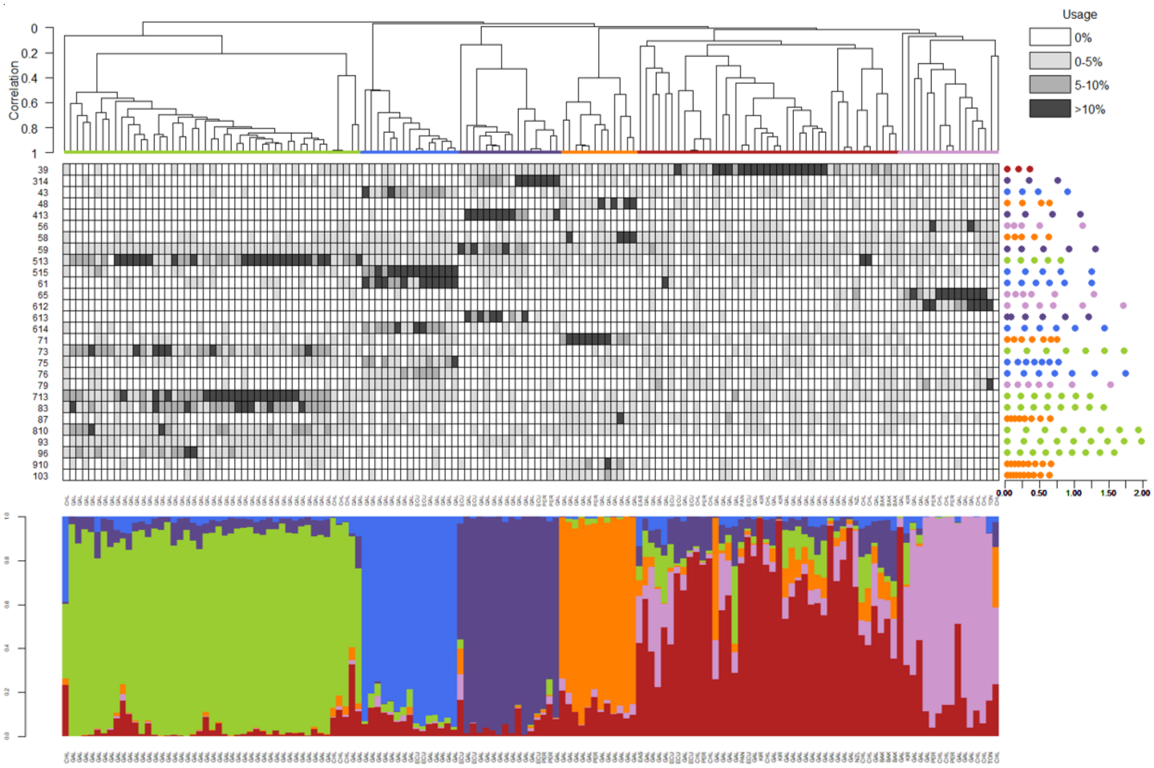
**D5.3:** *Figure S5.3* – Atlantic/Mediterranean tree with identity coda types. Average linkage hierarchical clustering dendrogram (top) depicts similarity among sperm whale coda repertoires (n=178) recorded in the Atlantic/Mediterranean. This dendrogram was constructed using *critfact*=10 and *minrep*=10. Colored identity clades correspond to three clans: Mediterranean (teal), EC1 (purple), and EC2 (gold). Heat map (middle) depicts identity coda type usage (rows) for each repertoire (columns) in shades of grey. Usage is calculated based on probabilistic assignment of codas to types (see Hersh et al. (2021) for details). Identity coda type numeric codes are on the left (see Table S5.4 for type names) and type centroid rhythm plots (colored by clan) are on the right of the heat map (each dot represents a click; scale bar is in seconds). Stacked bar plot (bottom) shows the posterior probabilities of repertoire assignations to clans (Whitehead & Hersh, submitted). Recording locations are listed along the bottom (see Table S5.1 for abbreviations).



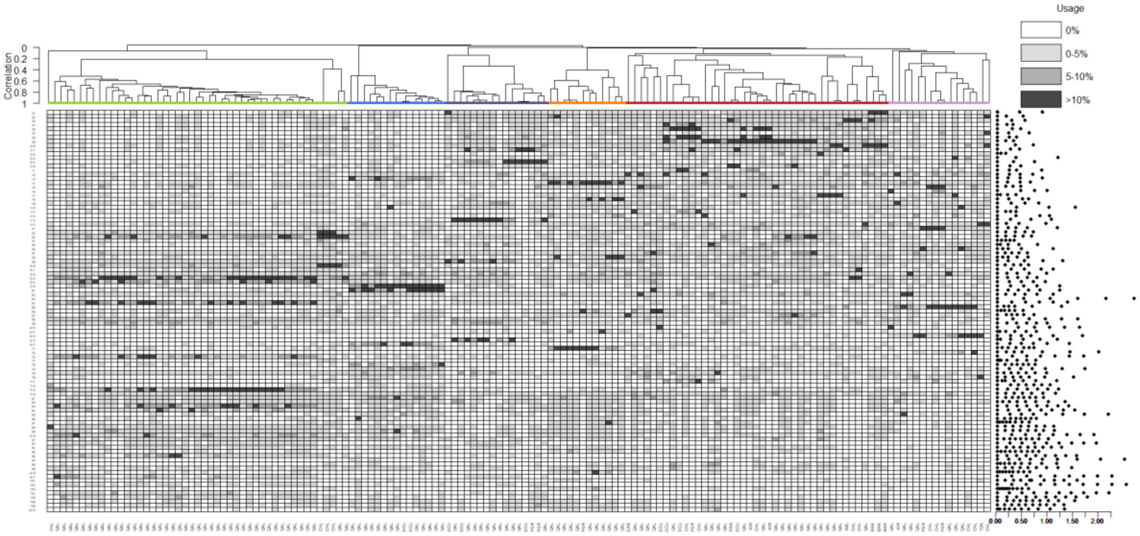
**D5.4:** *Figure S5.4* – Atlantic/Mediterranean tree with all (identity and non-identity) coda types. See Figure S5.3 for additional details and Table S5.4 for coda type names.



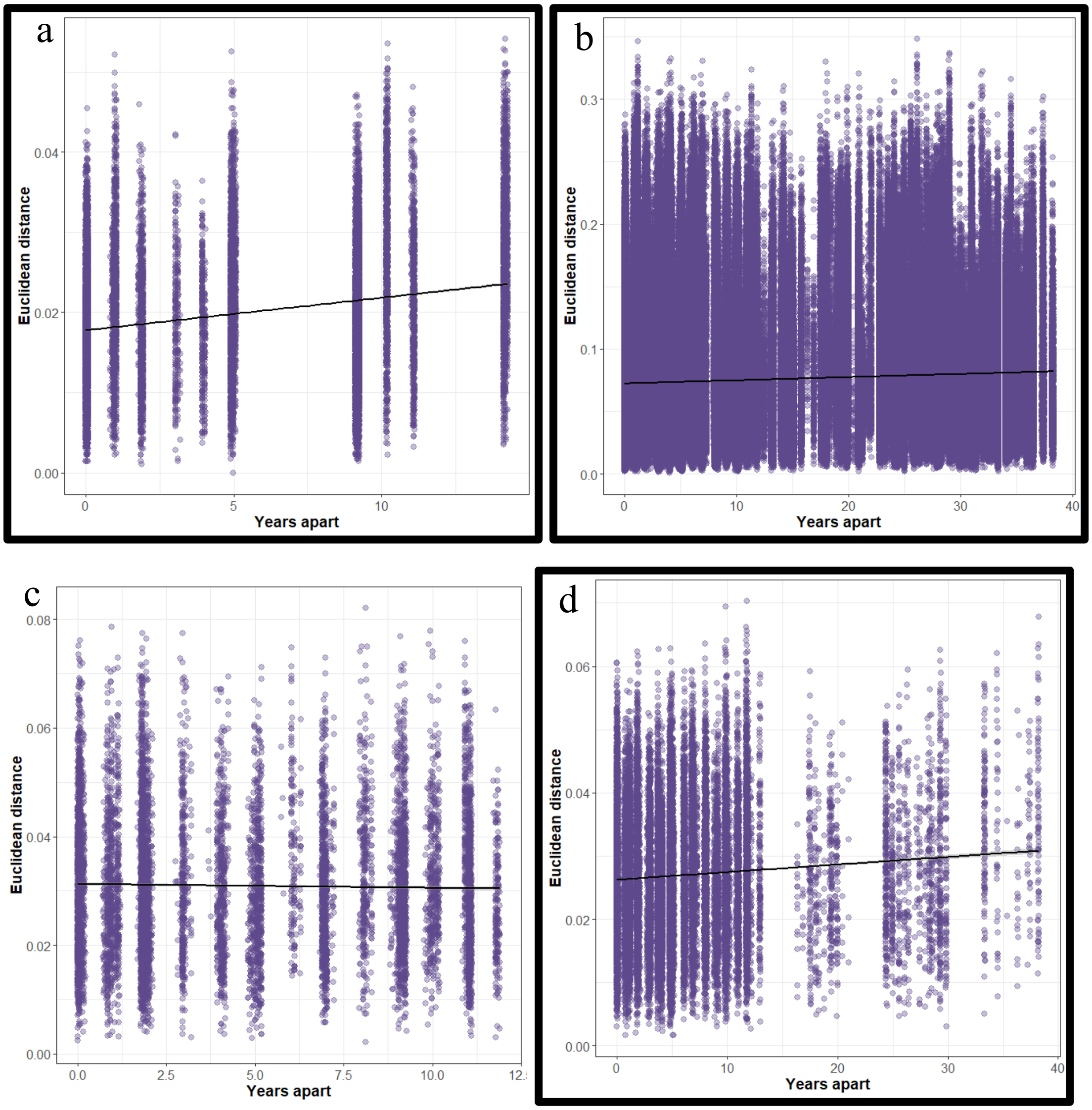
**D5.5:** *Figure S5.5* – Pacific tree with identity coda types. Average linkage hierarchical clustering dendrogram (top) depicts similarity among sperm whale coda repertoires (n=147) recorded in the Pacific. This dendrogram was constructed using *critfact*=10 and *minrep*=10. Colored identity clades correspond to six clans: Regular (green), Plus-One (blue), Slow Increasing (purple), Palindrome (orange), Short (red), and Four-Plus (pink). Heat map (middle) depicts identity coda type usage (rows) for each repertoire (columns) in shades of grey. Usage is calculated based on probabilistic assignment of codas to types (see Hersh et al. (2021) for details). Identity coda type numeric codes are on the left (see Table S5.5 for type names) and type centroid rhythm plots (colored by clan) are on the right of the heat map (each dot represents a click; scale bar is in seconds). Stacked bar plot (bottom) shows the posterior probabilities of repertoire assignations to clans (Whitehead & Hersh, submitted). Recording locations are listed along the bottom (see Table S5.1 for abbreviations).

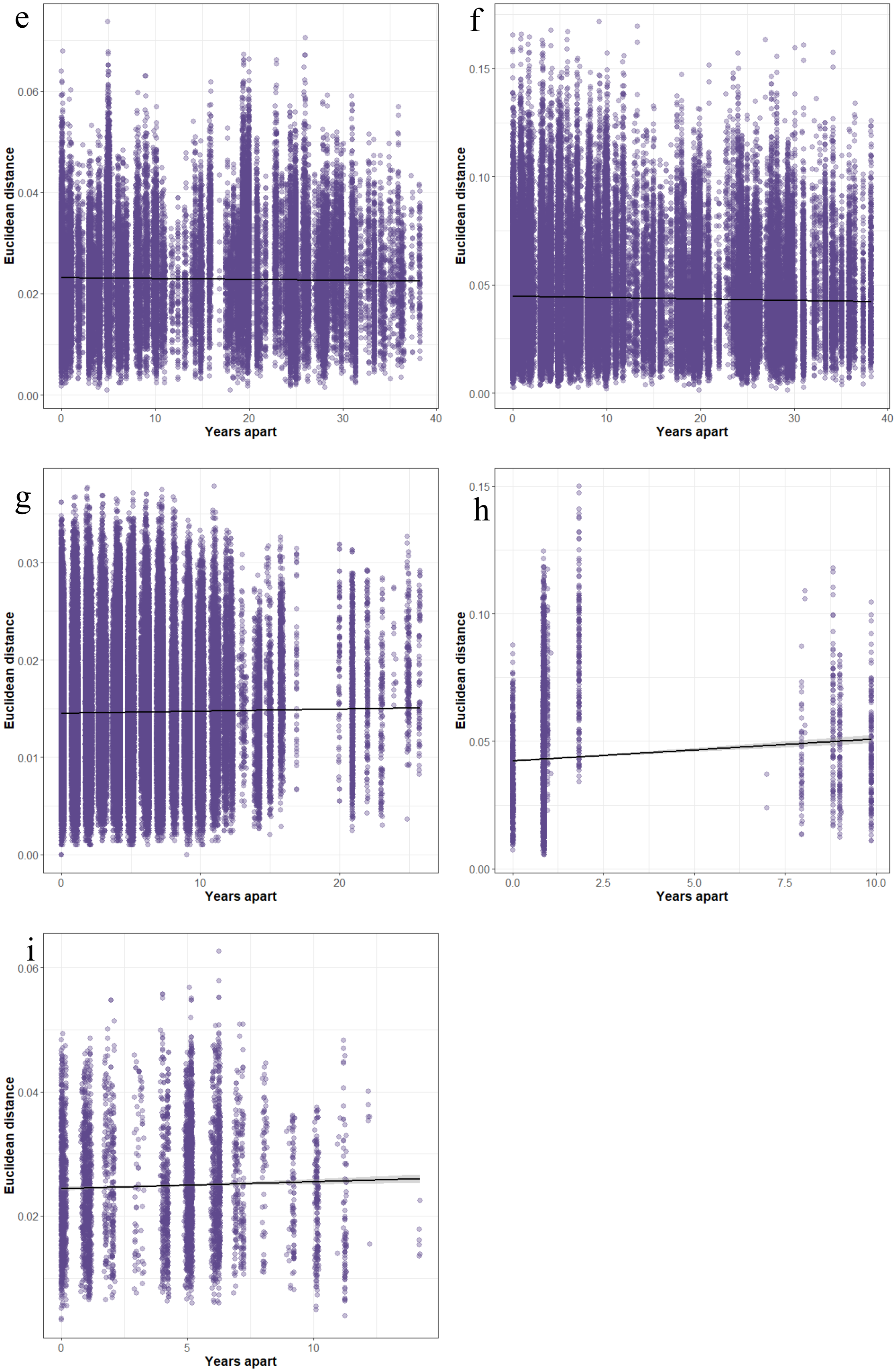


**D5.6:** *Figure S5.6* – Pacific tree with all (identity and non-identity) coda types. See Figure S5.5 for additional details and Table S5.5 for coda type names.

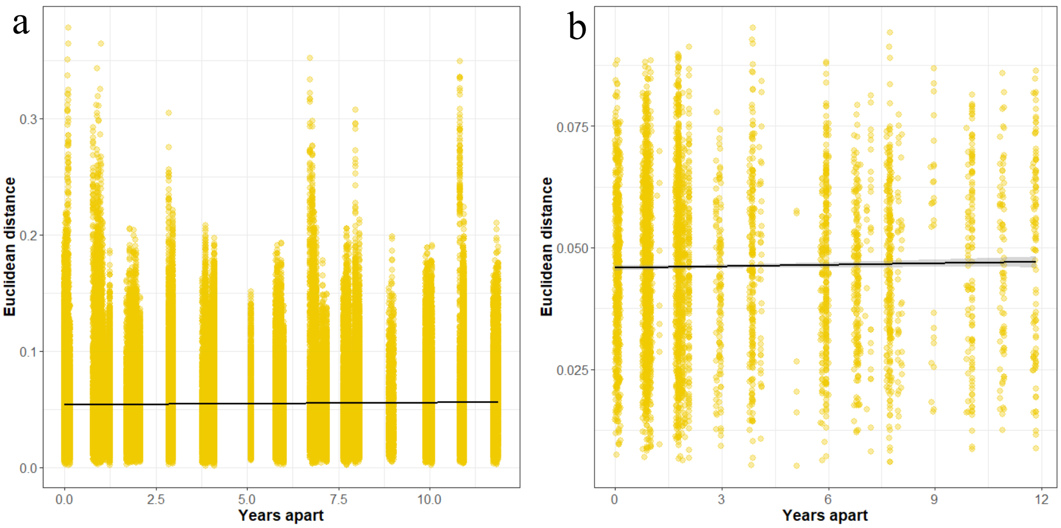


**D5.7:** *Figure S5.7* – Euclidean distance plots for nine EC1 clan coda types: (a) type 46, (b) type 51, (c) type 52, (d) type 54, (e) type 55, (f) type 57, (g) type 511, (h) type 69, and (i) type 84. Each point represents a pair of codas, with the number of years between coda recordings on the x-axis and the Euclidean distance between coda ICI vectors on the y-axis. Boxes indicate significant Mantel test results (see Table 5.1 for p-values), and the 95% confidence intervals are shown in gray.

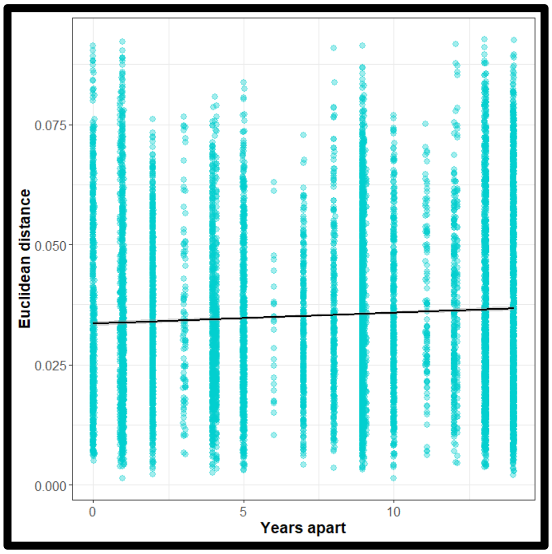



e
Euclidean distance
0.00
0.02
0.04
0.06
Years apart
0
10
20
30
40
f
Euclidean distance
0.00
0.05
0.10
0.15
Years apart
0
10
20
30
40
g
Euclidean distance
0.00
0.01
0.02
0.03
Years apart
0
10
20
h
Euclidean distance
0.00
0.05
0.10
0.15
Years apart
0.0
2.5
5.0
7.5
10.0
i
Euclidean distance
0.02
0.04
0.06
Years apart
0
5
10

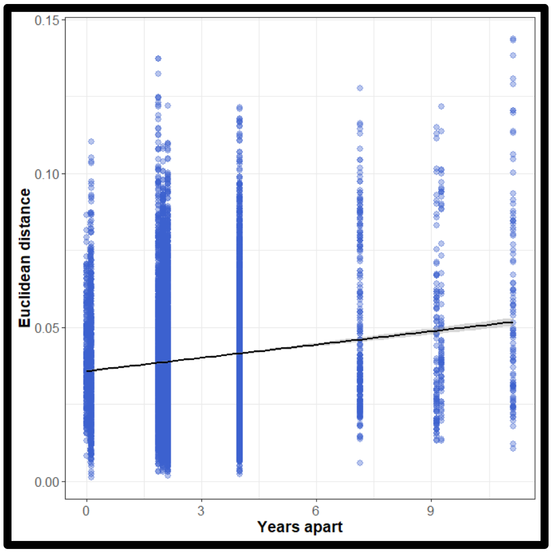
**D5.8:** *Figure S5.8* – Euclidean distance plots for two EC2 clan coda types: (a) type 514 and (b) type 515. Each point represents a pair of codas, with the number of years between coda recordings on the x-axis and the Euclidean distance between coda ICI vectors on the y-axis. Neither coda type had a significant Mantel test result (see Table 5.1 for p-values). The 95% confidence intervals are shown in gray.



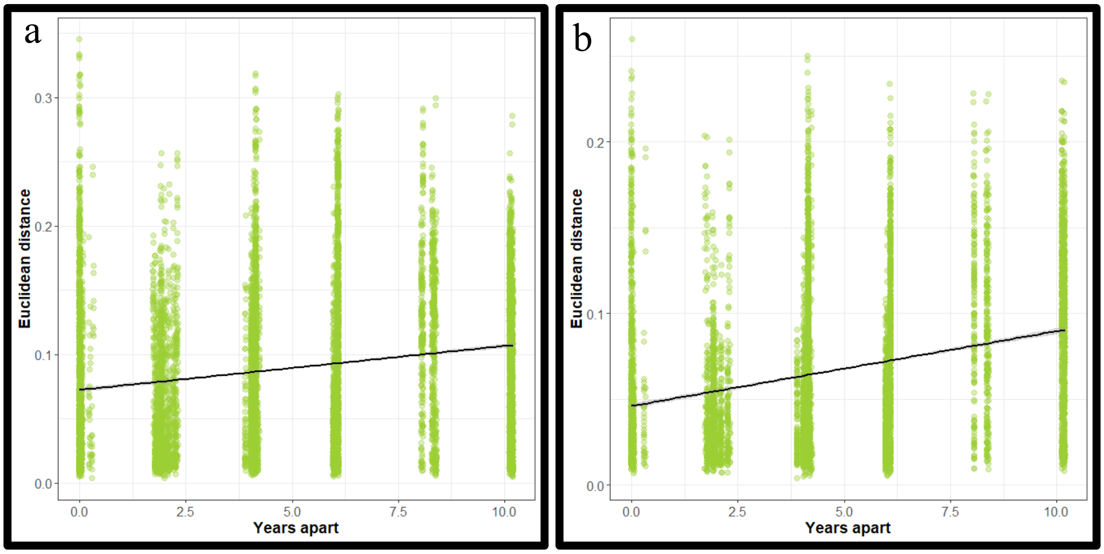
**D5.9:** *Figure S5.9* – Euclidean distance plot for Mediterranean clan coda type 412. Each point represents a pair of codas, with the number of years between coda recordings on the x-axis and the Euclidean distance between coda ICI vectors on the y-axis. The box indicates a significant Mantel test result (see Table 5.1 for p-value). The 95% confidence interval is shown in gray.

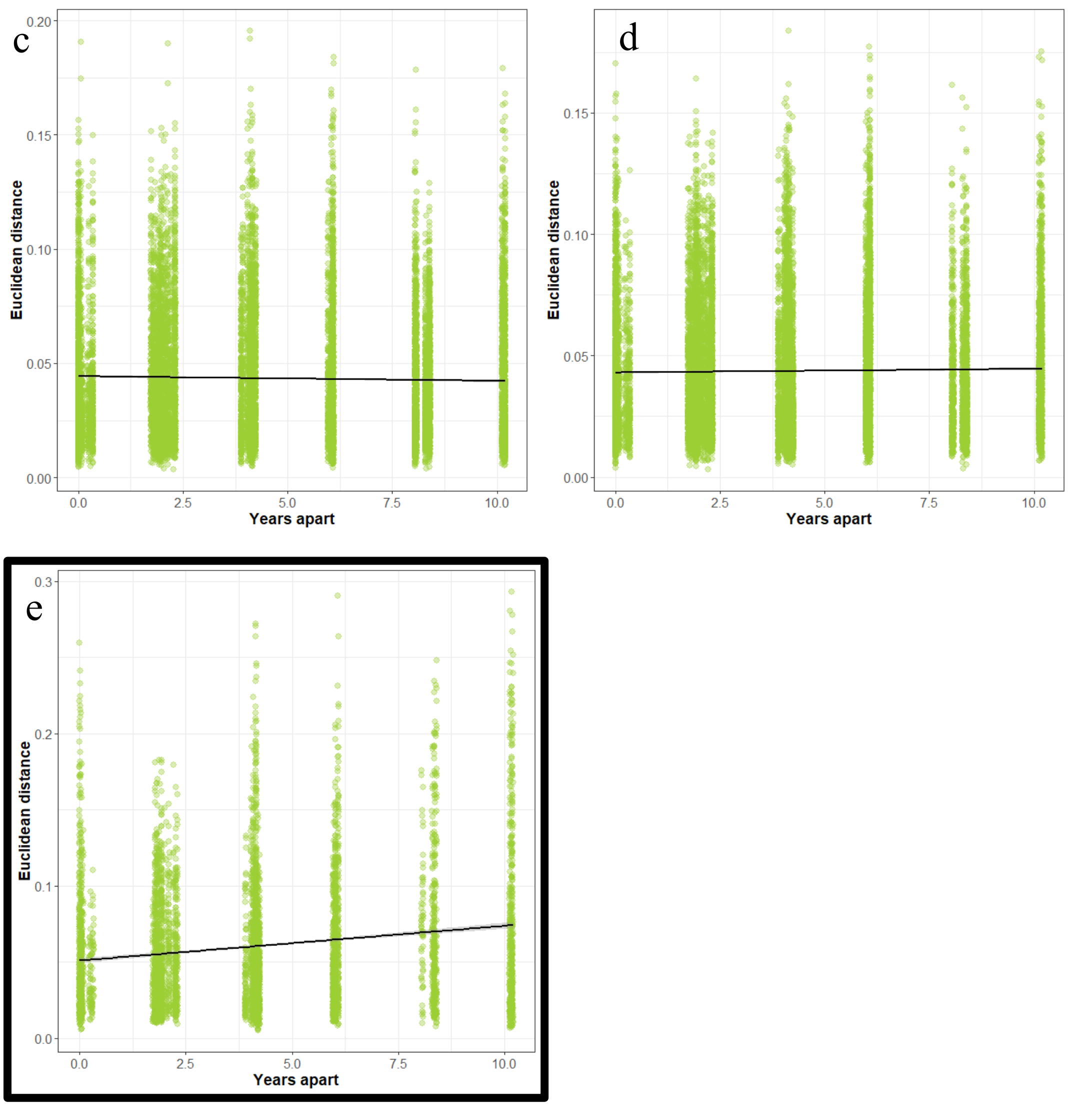


**D5.10:** *Figure S5.10* – Euclidean distance plot for Plus-One clan coda type 515. Each point represents a pair of codas, with the number of years between coda recordings on the x-axis and the Euclidean distance between coda ICI vectors on the y-axis. The box indicates a significant Mantel test result (see Table 5.1 for p-value). The 95% confidence interval is shown in gray.

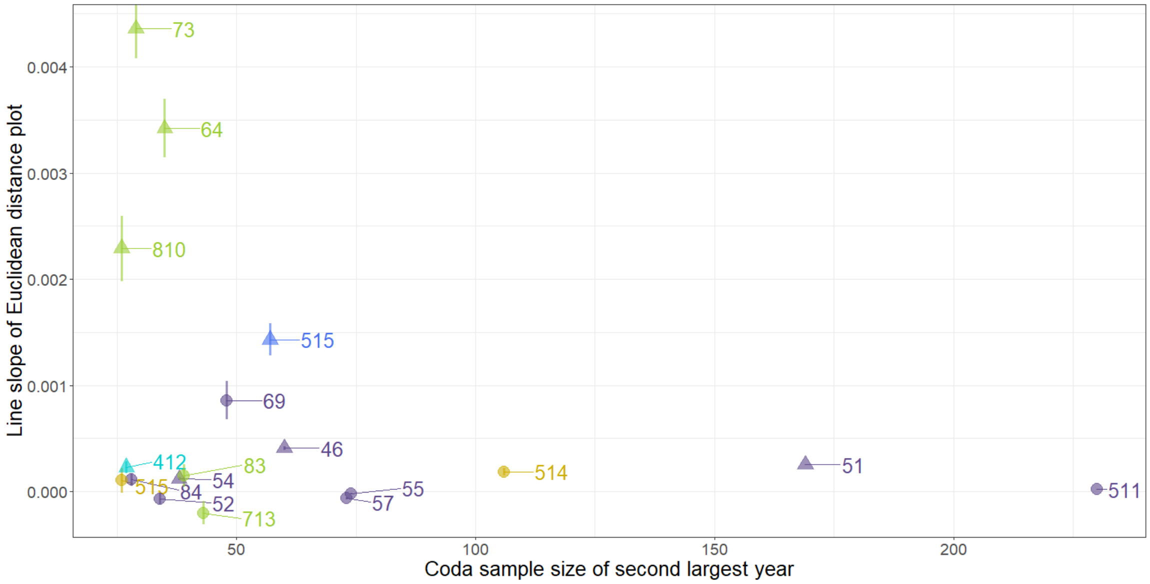


**D5.11:** *Figure S5.11* – Euclidean distance plots for five Regular clan coda types: (a) type 64, (b) type 73, (c) type 713, (d) type 83, and (e) type 810. Each point represents a pair of codas, with the number of years between coda recordings on the x-axis and the Euclidean distance between coda ICI vectors on the y-axis. Boxes indicate significant Mantel test results (see Table 5.1 for p-values), and the 95% confidence intervals are shown in gray.

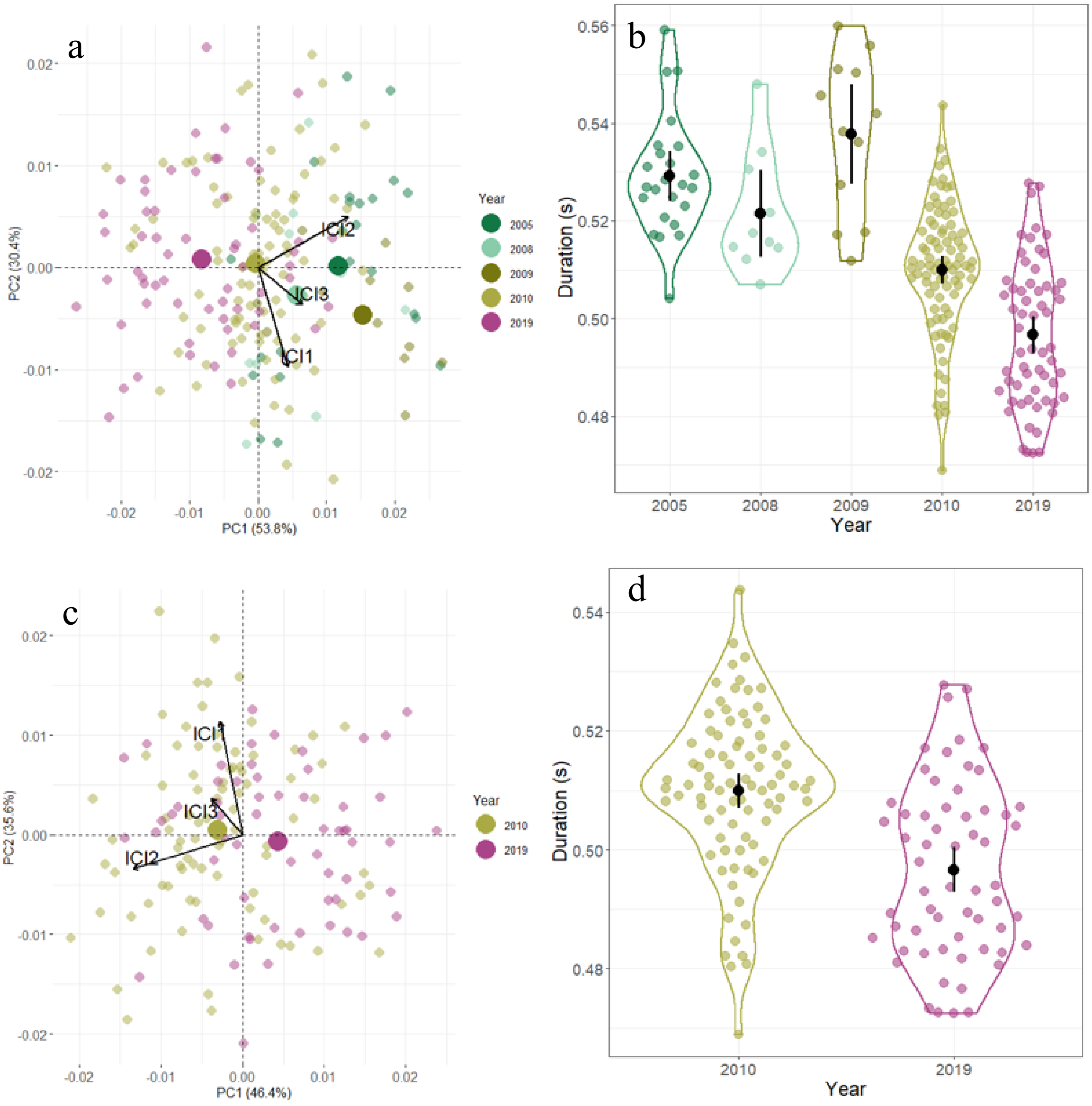



c
d
e
Euclidean distance
Years apart
0.00
0.05
0.10
0.15
0.20
0.0
0.1
0.2
0.3
0.0
2.5
5.0
7.5
10.0

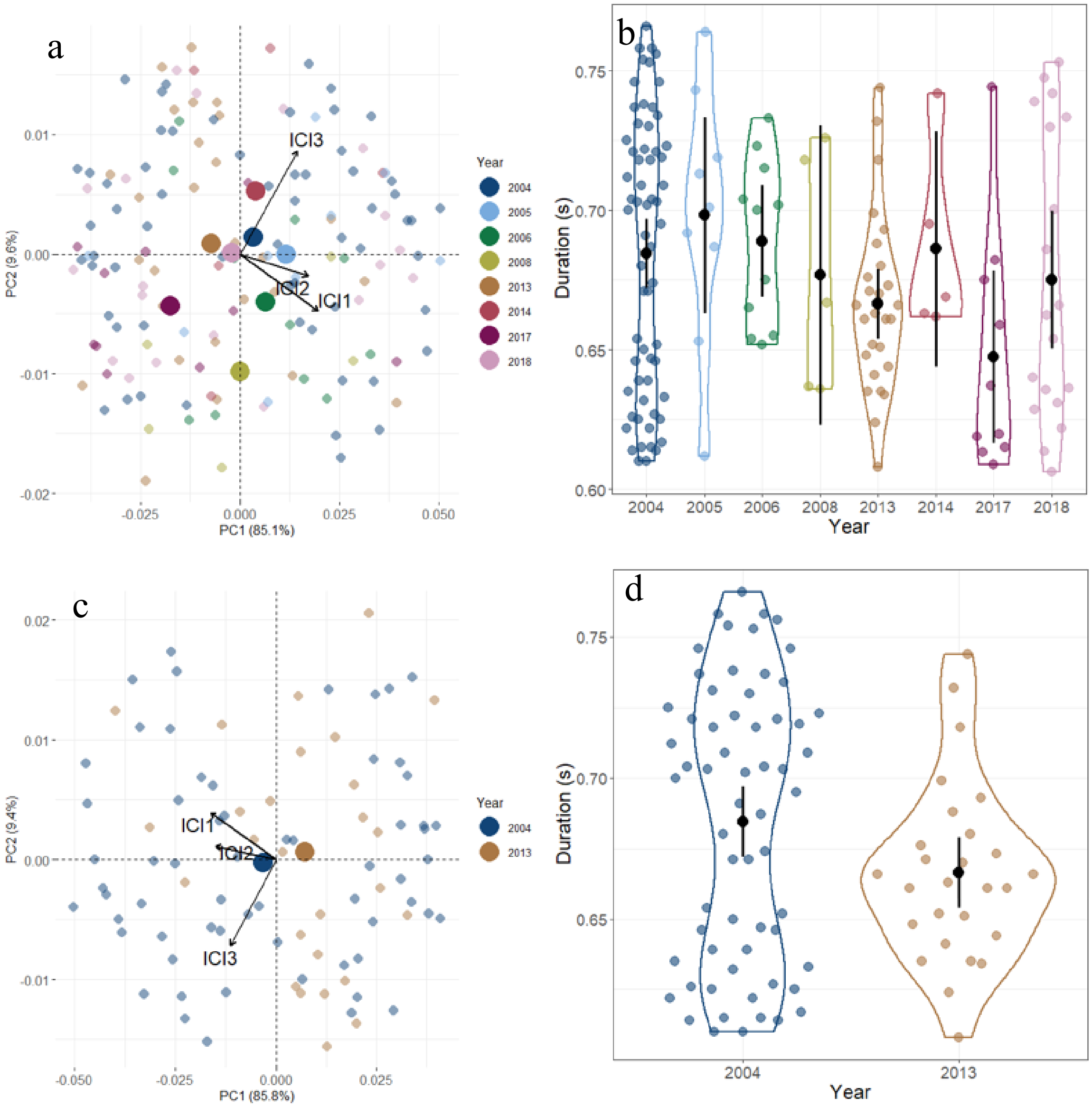
**D5.12:** *Figure S5.12* – Comparing the coda sample size of the second largest year to the slope (in s/yr) of the Euclidean distance plot. Each numbered point (see Tables S5.4/S5.6 for corresponding type names) represents the average slope of a different coda type. Bars denote 95% confidence intervals and colors correspond to clans: EC1=purple, EC2=gold, Mediterranean=teal, Plus-One=dark blue, Regular=green. Triangles indicate that the Mantel test result was significant, while circles indicate that it was not.



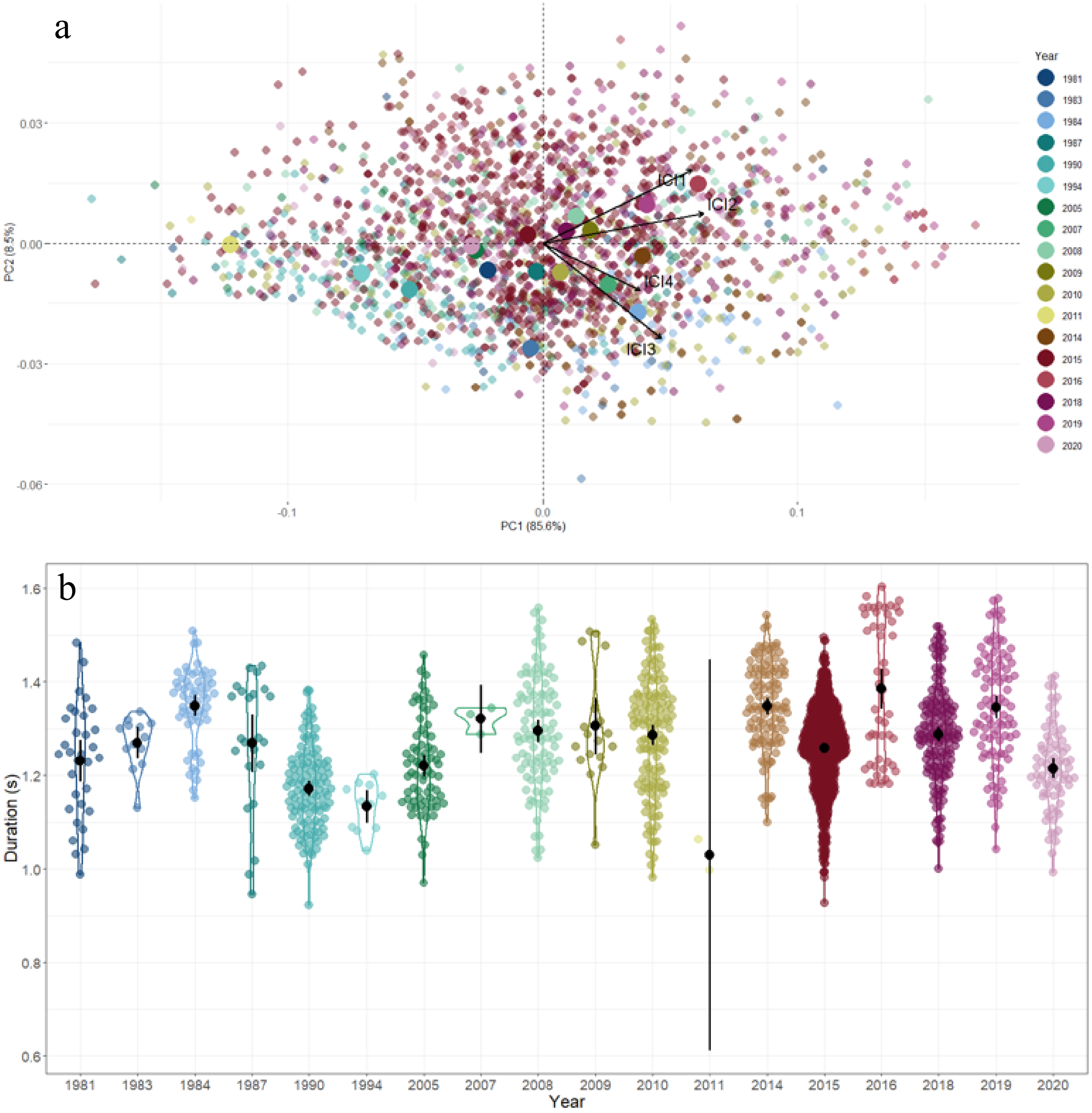
**D5.13:** *Figure S5.13* – EC1 clan coda type 46. Colors correspond to different years. Panels show the first two principal components for the (a) $\text{clan}_{\text{allyears}}$ and (c) $\text{clan}_{\text{largeyears}}$ analyses, as well as violin plots of coda duration with the mean and confidence limits in black for the (b) $\text{clan}_{\text{allyears}}$ and (d) $\text{clan}_{\text{largeyears}}$ analyses.

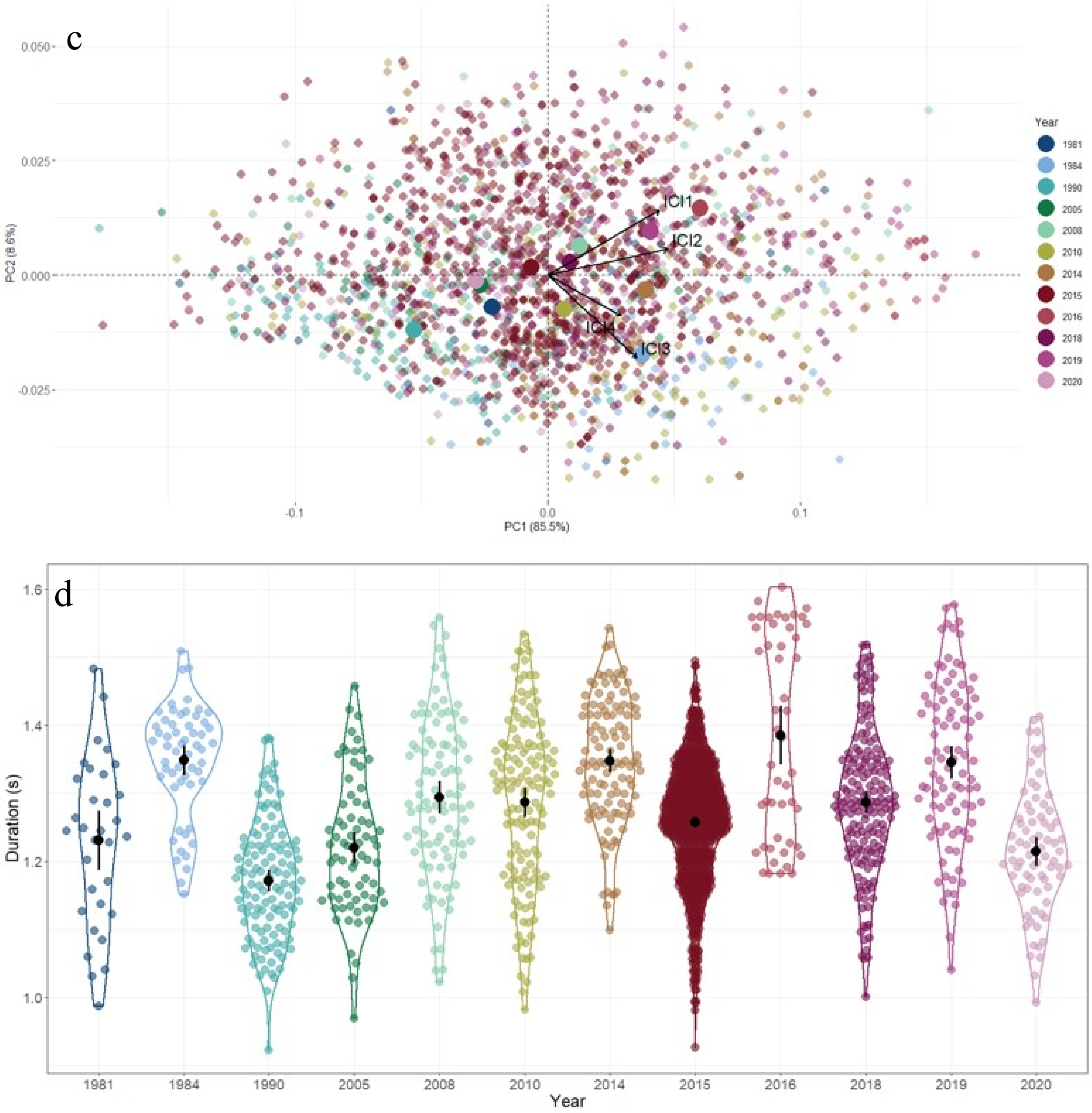


**D5.14:** *Figure S5.14* – Mediterranean clan coda type 412. Colors correspond to different years. Panels show the first two principal components for the (a) clan$_{\text{allyears}}$ and (c) clan$_{\text{largeyears}}$ analyses, as well as violin plots of coda duration with the mean and confidence limits in black for the (b) clan$_{\text{allyears}}$ and (d) clan$_{\text{largeyears}}$ analyses.

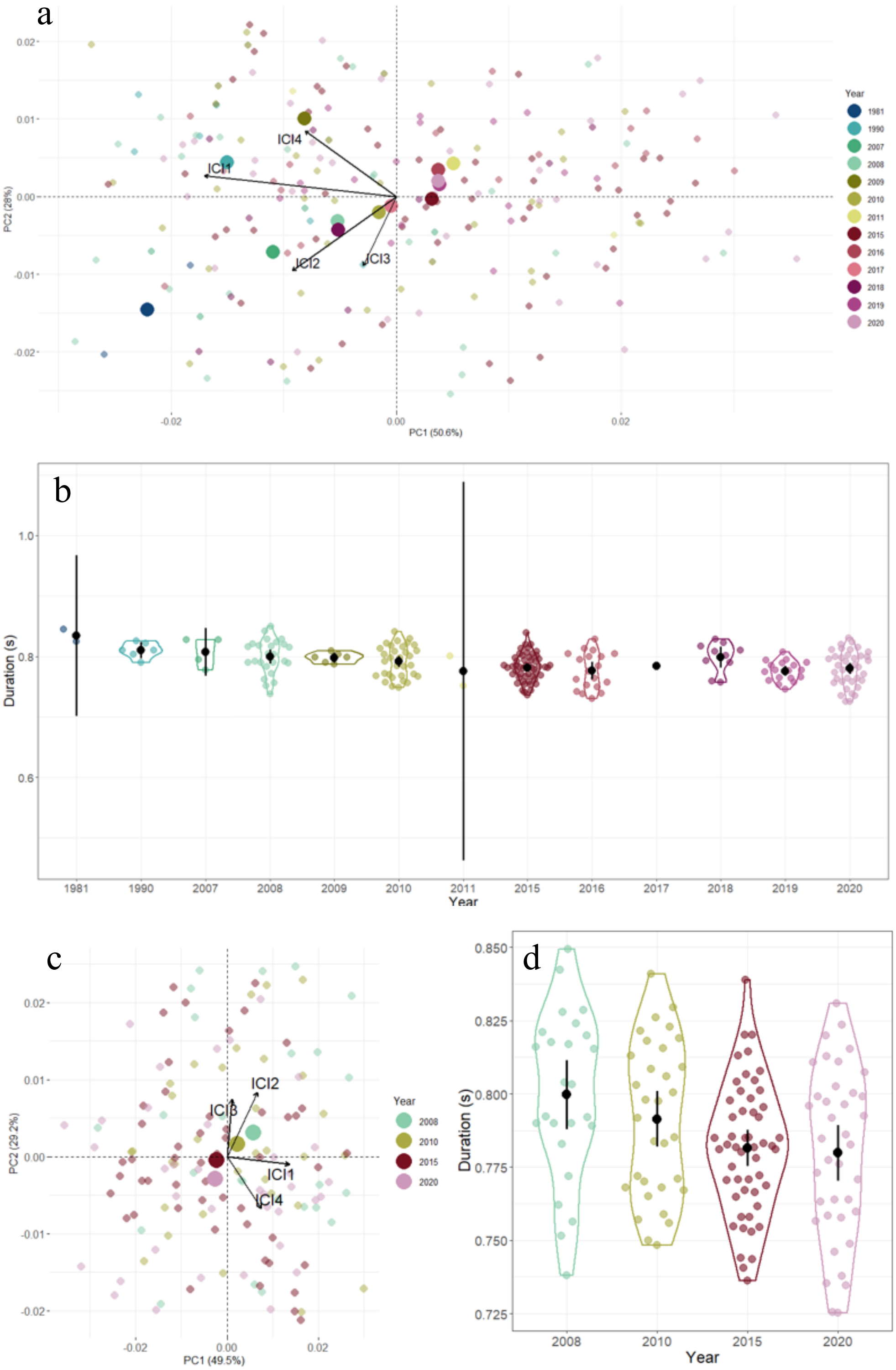


**D5.15:** *Figure S5.15* – EC1 clan coda type 51. Colors correspond to different years. Panels show the first two principal components for the (a) clan$_{\text{allyears}}$ and (c) clan$_{\text{largeyears}}$ analyses, as well as violin plots of coda duration with the mean and confidence limits in black for the (b) clan$_{\text{allyears}}$ and (d) clan$_{\text{largeyears}}$ analyses.

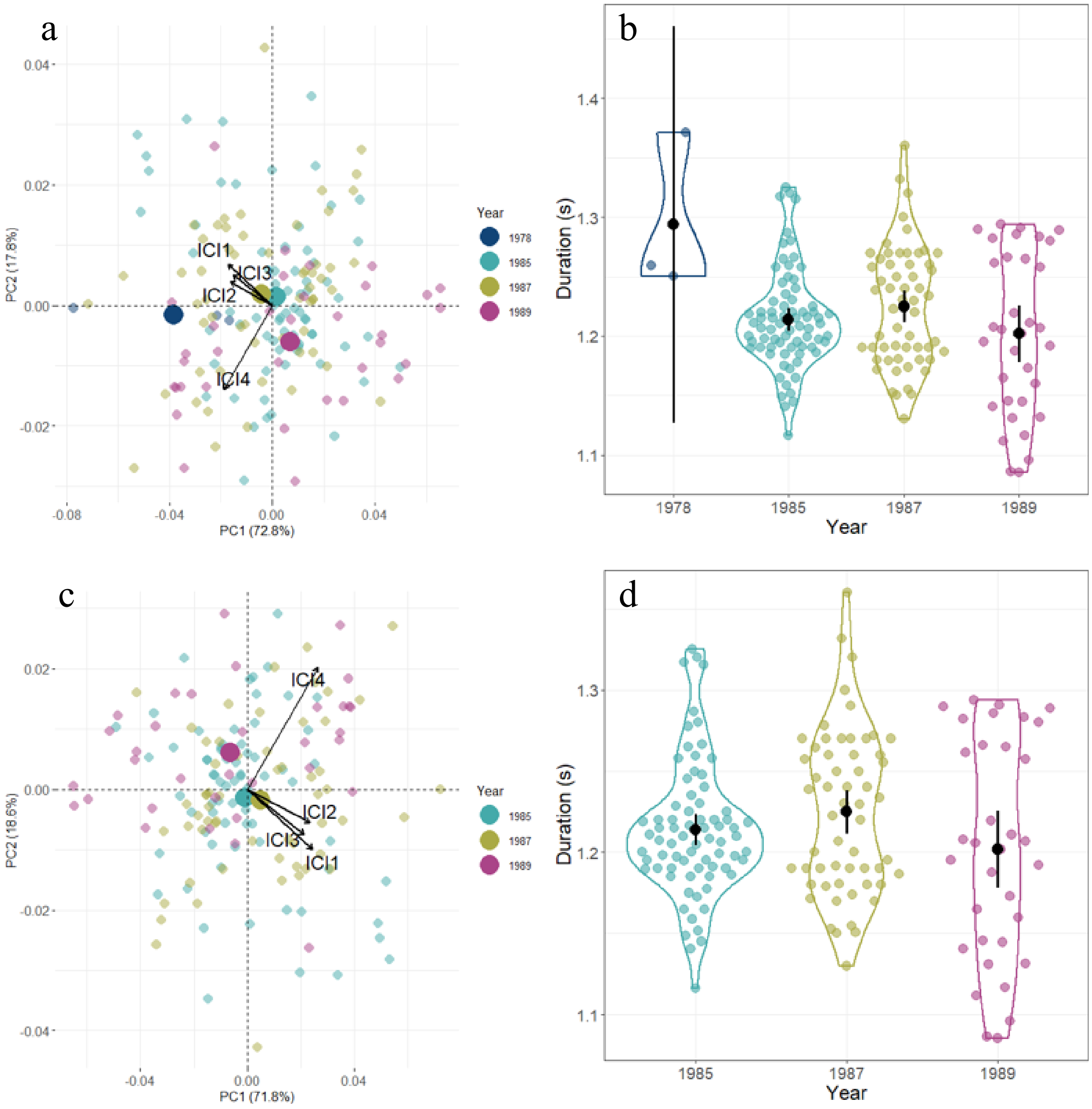



c
0.050
0.025
0.000
-0.025
PC2 (8.6%)
-0.1
0.0
0.1
PC1 (85.5%)
ICI1
ICI2
ICI3
ICI4
Year
1981
1984
1990
2005
2008
2010
2014
2015
2016
2018
2019
2020
d
1.6
1.4
1.2
1.0
Duration (s)
1981
1984
1990
2005
2008
2010
2014
2015
2016
2018
2019
2020
Year

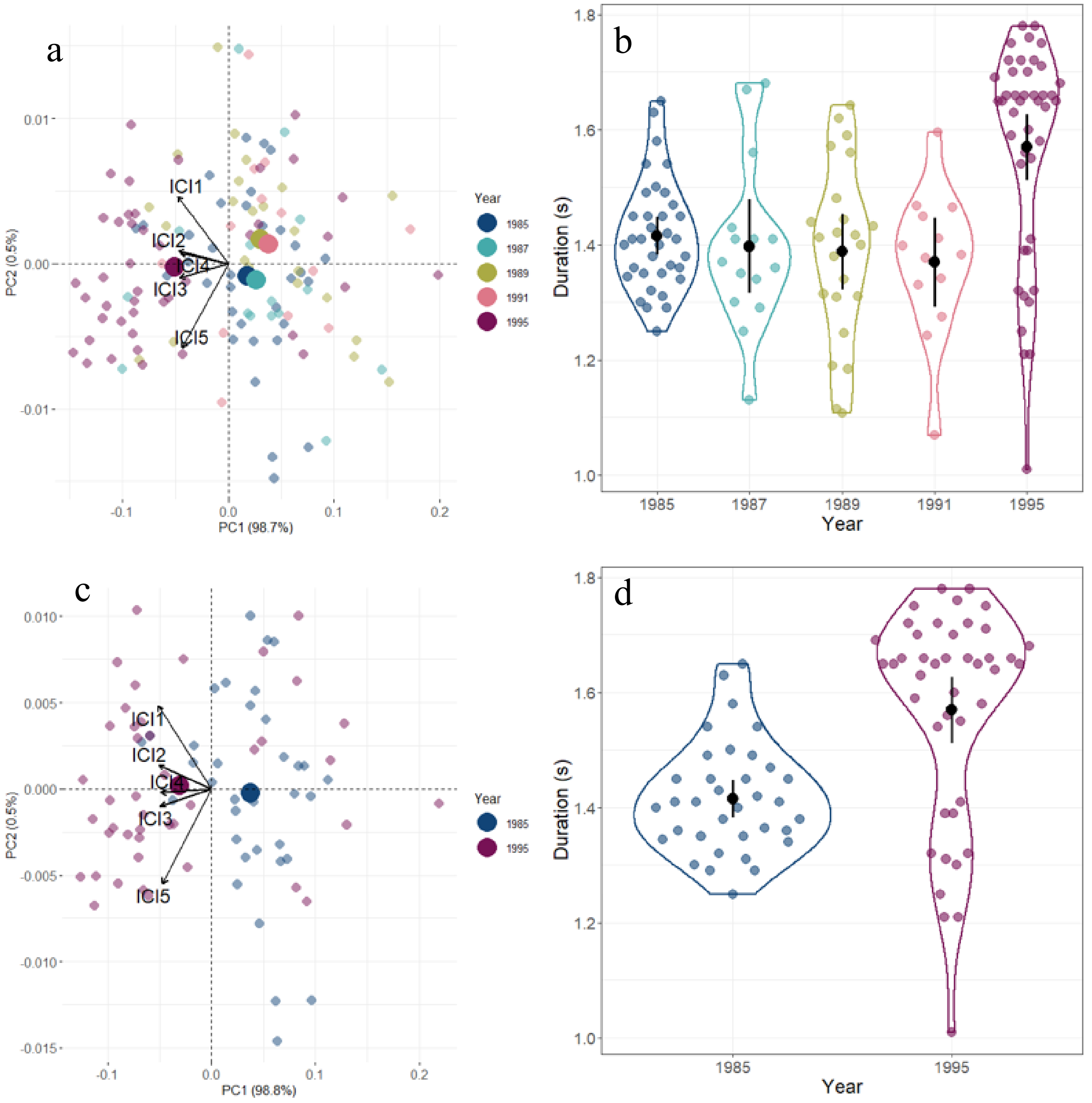
**D5.16:** *Figure S5.16* – EC1 clan coda type 54. Colors correspond to different years. Panels show the first two principal components for the (a) clan$_{allyears}$ and (c) clan$_{largeyears}$ analyses, as well as violin plots of coda duration with the mean and confidence limits in black for the (b) clan$_{allyears}$ and (d) clan$_{largeyears}$ analyses.



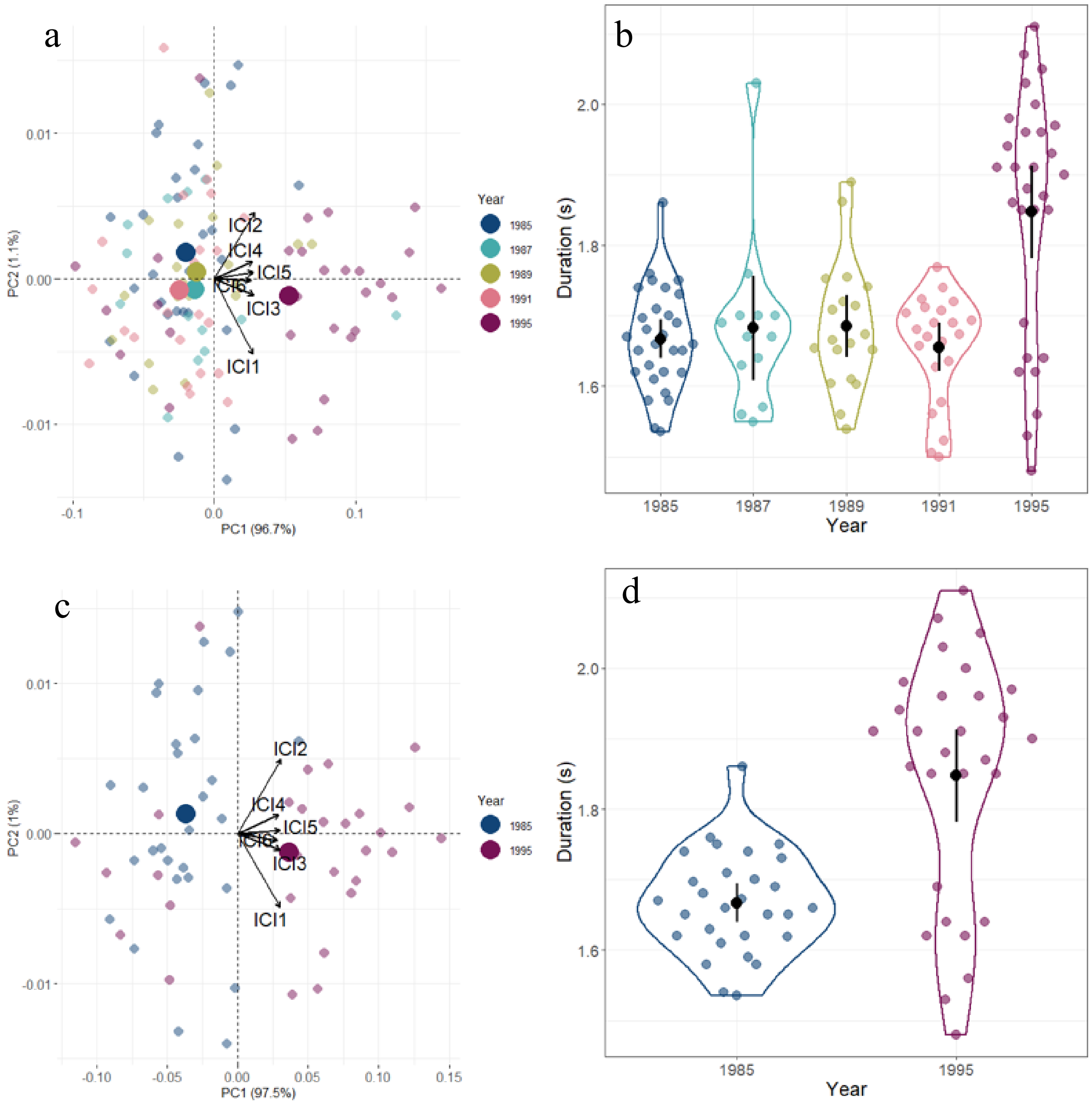
**D5.17:** *Figure S5.17* – Plus-One clan coda type 515. Colors correspond to different years. Panels show the first two principal components for the (a) clan$_{\text{allyears}}$ and (c) clan$_{\text{largeyears}}$ analyses, as well as violin plots of coda duration with the mean and confidence limits in black for the (b) clan$_{\text{allyears}}$ and (d) clan$_{\text{largeyears}}$ analyses.



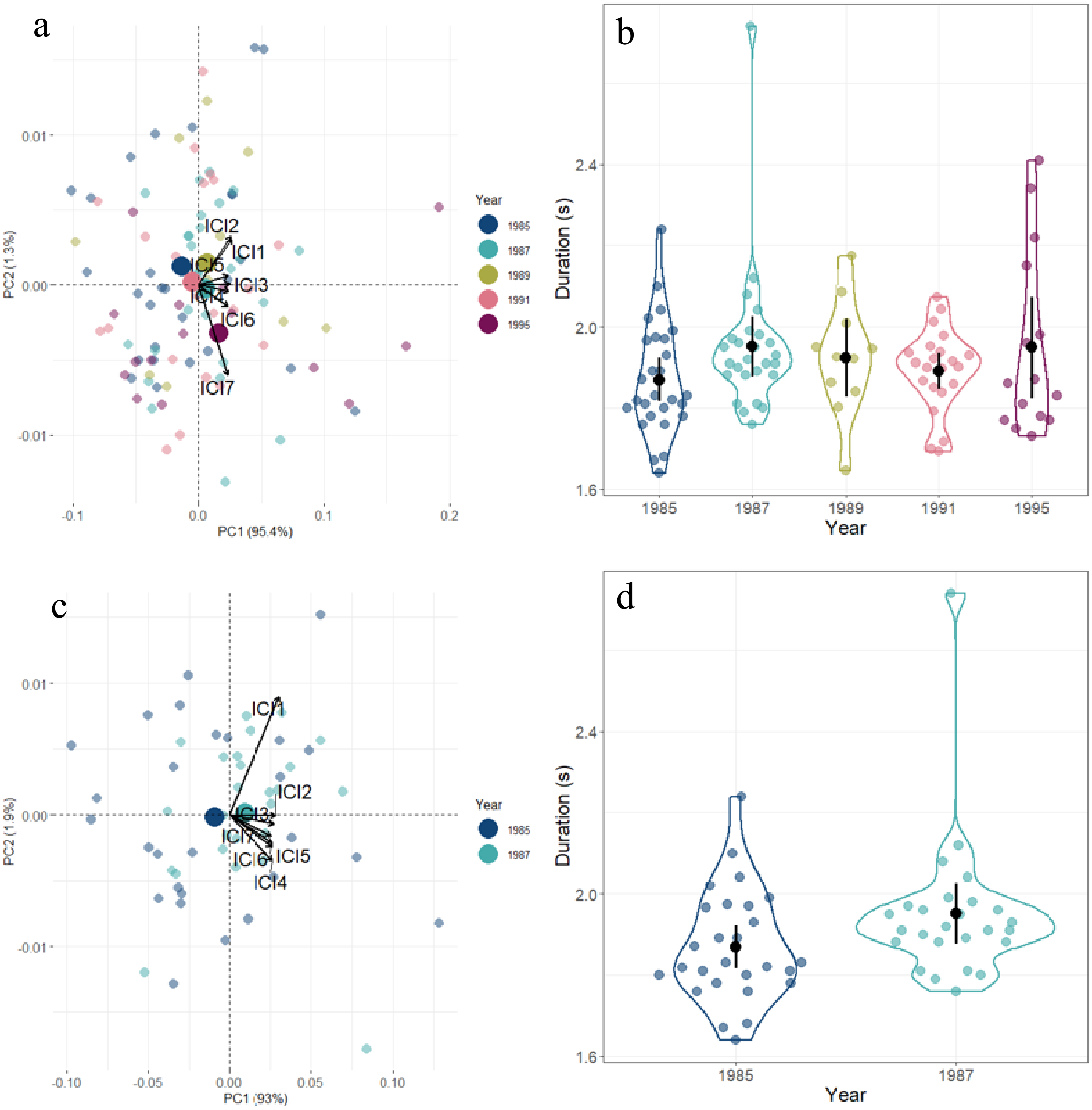
**D5.18:** *Figure S5.18* – Regular clan coda type 64. Colors correspond to different years. Panels show the first two principal components for the (a) $\text{clan}_{\text{allyears}}$ and (c) $\text{clan}_{\text{largeyears}}$ analyses, as well as violin plots of coda duration with the mean and confidence limits in black for the (b) $\text{clan}_{\text{allyears}}$ and (d) $\text{clan}_{\text{largeyears}}$ analyses.



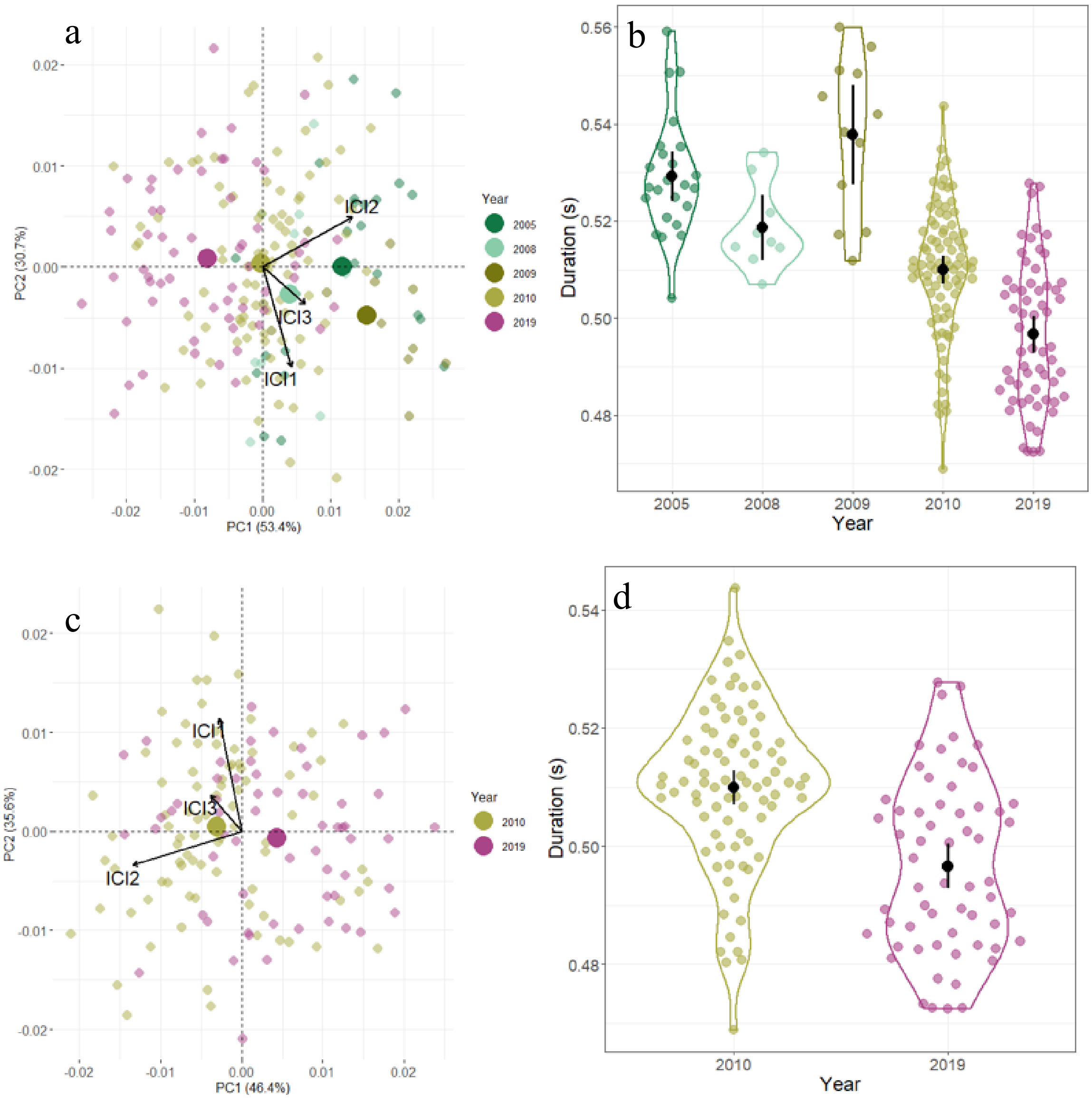
**D5.19:** *Figure S5.19* – Regular clan coda type 73. Colors correspond to different years. Panels show the first two principal components for the (a) clan$_{\text{allyears}}$ and (c) clan$_{\text{largeyears}}$ analyses, as well as violin plots of coda duration with the mean and confidence limits in black for the (b) clan$_{\text{allyears}}$ and (d) clan$_{\text{largeyears}}$ analyses.



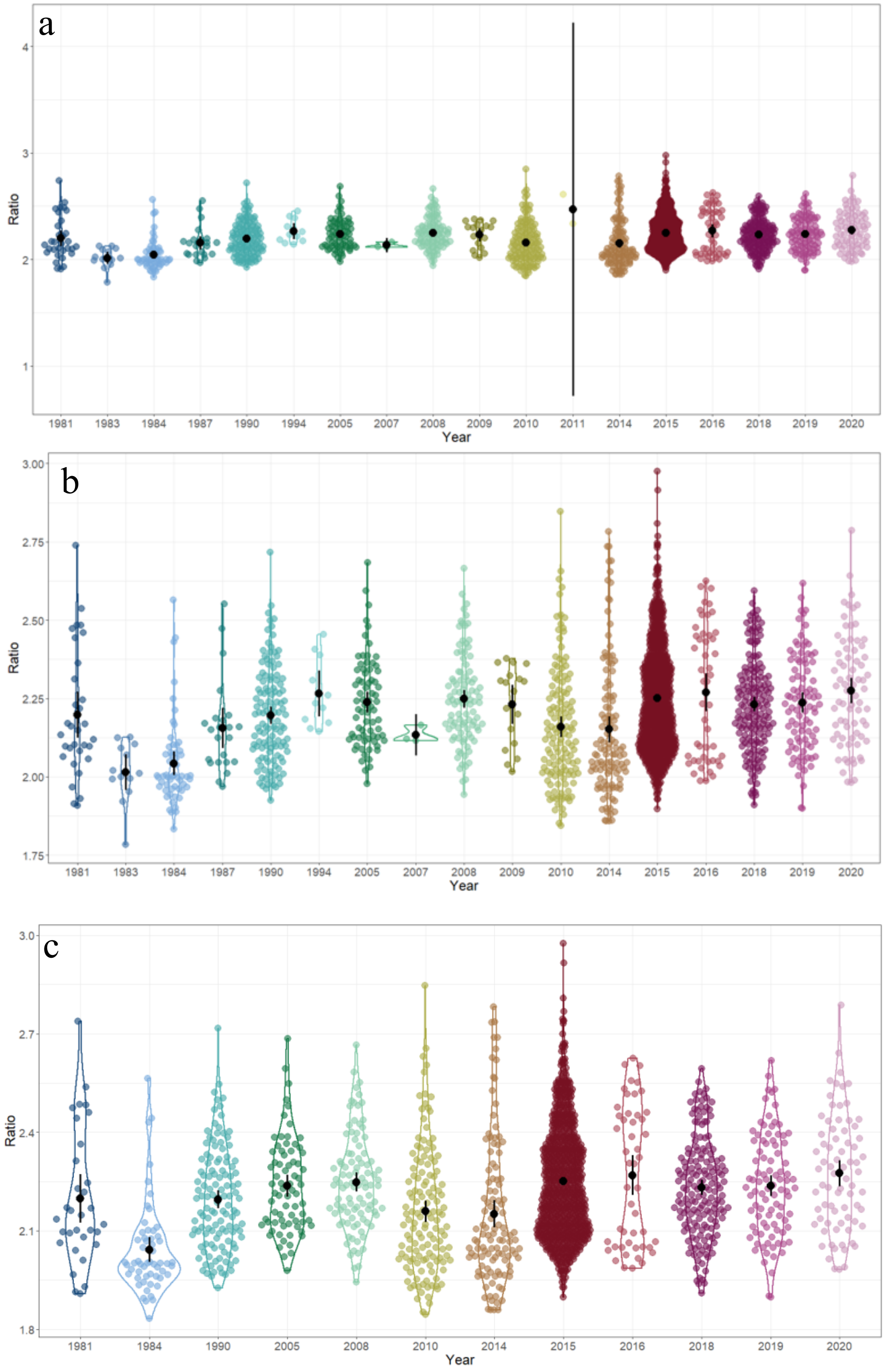
**D5.20:** *Figure S5.20* – Regular clan coda type 810. Colors correspond to different years. Panels show the first two principal components for the (a) $\text{clan}_{\text{allyears}}$ and (c) $\text{clan}_{\text{largeyears}}$ analyses, as well as violin plots of coda duration with the mean and confidence limits in black for the (b) $\text{clan}_{\text{allyears}}$ and (d) $\text{clan}_{\text{largeyears}}$ analyses.



**D5.21:** *Figure S5.21* – Social unit FU (EC1 clan) coda type 46. Colors correspond to different years. Panels show the first two principal components for the (a) clan$_{\text{allyears}}$ and (c) clan$_{\text{largeyears}}$ analyses, as well as violin plots of coda duration with the mean and confidence limits in black for the (b) clan$_{\text{allyears}}$ and (d) clan$_{\text{largeyears}}$ analyses.



**D5.22:** *Figure S5.22* – Ratio of long to short ICIs for EC1 clan coda type 51 over time. Colors correspond to different years. Panels show ratio violin plots with the mean and confidence limits in black for the (a) clan$_{allyears}$ (n=1,972 codas) and (c) clan$_{largeyears}$ (n=1,902 codas) analyses. Panel (b) is the same as panel (a) but with year 2011 removed to allow for better visualization. For both analyses, the fixed effect (i.e. year) model coefficient was 0.003 and the p-values were 0.002 and 0.012, respectively.



**D5.31:** *Table S5.1* – Complete sperm whale coda dataset used in IDcall. The 'Lesser Antilles' include Antigua & Barbuda, Dominica, Guadeloupe, Grenada, St. Kitts & Nevis, St. Lucia, Montserrat, Martinique, and St. Vincent & the Grenadines. 'Kiribati' was referred to as 'Jarvis Island' in Hersh et al. (2021). Gray shading denotes regions included in the Pacific tree; no shading denotes regions included in the Atlantic/Mediterranean tree.

| **Recording location** | **Abbreviation** | **Number of codas (3–10 clicks)** | **Number of repertoires (≥25 codas)** |
|---|---|---|---|
| Lesser Antilles | LAN | 18,121 | 159 |
| Panama | PAN | 353 | 1 |
| Gulf of Mexico | GOM | 102 | 1 |
| Balearic Islands | BAL | 1,967 | 17 |
| Panama | PAN | 191 | 1 |
| Galápagos Islands | GAL | 10,315 | 108 |
| Ecuador | ECU | 774 | 9 |
| Peru | PER | 665 | 6 |
| Chile | CHL | 5,706 | 14 |
| Easter Island | EAS | 90 | 1 |
| Kiribati | KIR | 511 | 3 |
| Baker Island | BAK | 272 | 3 |
| Tonga | TON | 152 | 1 |
| New Zealand | NZL | 44 | 1 |
| | **Total:** | **39,263** | **325** |

**D5.32:** *Table S5.2* – Summary of clans in the Atlantic/Mediterranean and Pacific dendrograms. For each repertoire, we calculated the correlation between the coda type usages of the repertoire and the median usages of the clan and averaged these values to get the within-clan correlation. Gray shading denotes regions included in the Pacific tree; no shading denotes regions included in the Atlantic/Mediterranean tree.

| **Clan name** | **Number of repertoires** | **Within-clan correlation (mean ± SD)** | **Number of identity codas** |
|---|---|---|---|
| EC1 | 121 | 0.547 ± 0.246 | 5 |
| EC2 | 38 | 0.652 ± 0.323 | 7 |
| Mediterranean | 17 | 0.833 ± 0.110 | 6 |
| Regular | 47 | 0.741 ± 0.233 | 7 |
| Plus-One | 15 | 0.842 ± 0.135 | 6 |
| Slow Increasing | 16 | 0.706 ± 0.208 | 4 |
| Palindrome | 12 | 0.725 ± 0.187 | 6 |
| Short | 41 | 0.456 ± 0.260 | 1 |
| Four-Plus | 16 | 0.504 ± 0.270 | 4 |

**D5.33:** *Table S5.3* – Codas included in within-type analyses of temporal stability. 'Recording year span' includes the first and last years with codas. The final four columns detail how the original clan coda datasets were sequentially restricted, and the last column gives the number of codas ultimately analyzed for each clan (see text for details). Gray shading denotes regions included in the Pacific tree; no shading denotes regions included in the Atlantic/Mediterranean tree. See Table S5.1 for recording location abbreviations.

| **Recording location** | **Clan** | **Recording year span** | **Recording years** | **Total codas** | **Codas after outlier removal** | **Codas after coda-to-type restriction** | **Codas after repertoire-to-clan restriction** |
|---|---|---|---|---|---|---|---|
| LAN | EC1 | 40 | 19: 81, 83, 84, 87, 90, 94, 95<br>20: 05, 07, 08, 09, 10, 11, 14, 15, 16, 17, 18, 19, 20 | 13,190 | 12,124 | 6,886 | 6,148 |
| | EC2 | 40 | 19:81<br>20: 08, 11, 12, 16, 18, 19, 20 | 4,931 | 4,397 | 1,991 | 1,437 |
| BAL | Mediterranean | 15 | 20: 04, 05, 06, 07, 08, 13, 14, 17, 18 | 1,750 | 1,553 | 641 | 574 |
| GAL | Plus-One | 12 | 19: 78, 85, 87, 89 | 1,770 | 1,308 | 1,058 | 404 |
| | Regular | 11 | 19: 85, 87, 89, 91, 95 | 3,694 | 2,907 | 1,688 | 1,102 |

**D5.34:** *Table S5.4* – Atlantic/Mediterranean numeric codes and names for all coda types. We give the number of clicks, the numeric code (see Figure S5.4), and the type name following conventions in Weilgart & Whitehead (1997) and Hersh et al. (2021) (see Table 4.1 for additional details). Bolded numeric codes denote identity codas and are colored by clan (see Figure S5.3 for colors).

| Number of clicks | Numeric code | Type name |
|---|---|---|
| 3 | 32, 34, 35 | 3R |
| | **31**, **36** | 2+1 |
| | 33 | 1+2 |
| 4 | 42, 47, 49, 411 | 4R |
| | 45 | 4I |
| | 46 | 4D |
| | **41**, **410**, **412**, **413** | 3+1 |
| | 48 | 1+3 |
| | 44 | 1+2+1 |
| 5 | **51**, **54**, **56**, **57**, **511**, 512 | 1+1+3 |
| | 55, 58, 59, 510, 513, **514**, **515** | 5R |
| | 52 | 5D |
| | 53 | 2+1+1+1 |
| 6 | 63, 64, **65**, 610 | 6R |
| | 61 | 6D |
| | 67 | 5+1 |
| | 69 | 1+5 |
| | 611, 612 | 1+1+4 |
| | 62 | 4+1++1 |
| | 68 | 1++1+4 |
| 7 | 74, 76, 78, **710**, 711 | 7R |
| | 75 | 7I |
| | 73, 77 | 1+6 |
| | 71 | 1+1+5 |
| | 72 | 5+1++1 |
| 8 | **81**, 86, 88, **811** | 8R |
| | 82 | 1+1+6 |
| | 84, 89 | 7+1 |
| | 87, 810 | 1+7 |
| 9 | 91, 93, 98, 99, 910, 914 | 9R |
| | 92, 95, 912 | 8+1 |
| | 94 | 1+8 |
| | 96 | 1+1+7 |
| 10 | 103, 107, **109** | 10R |
| | 101 | 10I |
| | 105, 106 | 9+1 |

**D5.36:** *Table S5.5* – Pacific numeric codes and names for all coda types. We give the number of clicks, the numeric code (see Figure S5.6), and the type name following conventions in Weilgart & Whitehead (1997) and Hersh et al. (2021) (see table 4.1 for additional details). Bolded numeric codes denote identity codas and are colored by clan (see Figure S5.5 for colors).

| **Number of clicks** | **Numeric code** | **Type name** |
|---|---|---|
| 3 | 31, **39**, 310 | 3R |
|  | 32, 37, 312, 313, **314** | 2+1 |
|  | 33, 34, 36, 38, 311, 315 | 1+2 |
| 4 | 42, 44, 45, 49 | 4R |
|  | 41 | 1+2++1 |
|  | **43** | 3+1 |
|  | 46, 411 | 1+3 |
|  | 47, **413**, 414 | 4I |
|  | **48**, 410 | 2+2 |
|  | 412 | 1+1+2 |
| 5 | 52, 53, 54, 57, 510, 511, **513**, 514 | 5R |
|  | 51, 512, **515** | 4+1 |
|  | 55, **59** | 5I |
|  | **58** | 3+1+1 |
|  | **56** | 3+1++1 |
| 6 | 63, 64, 68, 69, 610 | 6R |
|  | **61**, 62, **614** | 5+1 |
|  | **65**, **612** | 4+1++1 |
|  | 66, 611, **613** | 6I |
|  | 67 | 1+4+1 |
| 7 | **73**, **75**, **713**, 715 | 7R |
|  | **71** | 3+1+3 |
|  | 72, 74 | 4+1++1++1 |
|  | **76**, 78, **79** | 5+1++1 |
|  | 711 | 7I |
|  | 712 | 1+6 |
| 8 | 81, **83**, 84, 86, 88, 89, **810** | 8R |
|  | 82, 85, **87** | 8I |
| 9 | **93**, 94, **96**, 911 | 9R |
|  | 91 | 8+1 |
|  | 95, 97, 98, 99, **910** | 9I |
| 10 | 102, 104, 105, 109 | 10R |
|  | 101 | 4+2+4 |
|  | **103**, 1010 | 10I |
|  | 106 | 9+1 |

# APPENDIX E – COPYRIGHT PERMISSIONS

**E3.1:** *Copyright Release S3.1* – Permission to reproduce published article.

This Agreement between Dr. Taylor Hersh ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

| | |
|---|---|
| License Number | 5204291070552 |
| License date | Dec 08, 2021 |
| Licensed Content Publisher | John Wiley and Sons |
| Licensed Content Publication | Methods in Ecology & Evolution |
| Licensed Content Title | Using identity calls to detect structure in acoustic datasets |
| Licensed Content Author | Taylor A. Hersh, Shane Gero, Luke Rendell, et al |
| Licensed Content Date | May 29, 2021 |
| Licensed Content Volume | 12 |
| Licensed Content Issue | 9 |
| Licensed Content Pages | 11 |
| Type of use | Dissertation/Thesis |
| Requestor type | Author of this Wiley article |
| Format | Electronic |
| Portion | Full article |
| Will you be translating? | No |
| Title | Dialects over space and time: How codas augment sociality in sperm whales |
| Institution name | Dalhousie University |
| Expected presentation date | Dec 2021 |
| Requestor Location | Dr. Taylor Hersh<br>Platolaan 698<br>Nijmegen, 6525 KK<br>Netherlands<br>Attn: Dr. Taylor Hersh |

| | |
|---|---|
| Publisher Tax ID | EU826007151 |
| Total | 0.00 USD |