Leveraging eDNA to detect and monitor hybrid zones

Kathryn A. Stewart* and Scott A. Taylor

1Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 904 Science Park, 1098 XH Amsterdam, North Holland, NL.

2Department Ecology and Evolutionary Biology, University of Colorado Boulder, 1900 Pleasant Street, 80309 Boulder CO, USA.

* Corresponding Author: Kathryn A. Stewart, K.stewart@uva.nl

Abstract

Hybrid zones are important windows into evolutionary processes and our understanding of their significance and prevalence in nature has expanded quickly. Yet most hybridization research has restricted temporal and spatial resolution, limiting our ability to draw broad conclusions about evolutionary and conservation related outcomes. Here, we argue rapidly advancing environmental DNA (eDNA) methodology should be adopted for studies of hybrid zones to increase temporal sampling (contemporary and historical), refine and geographically expand sampling density, and collect data for taxa that are difficult to directly sample. Genomic data in the environment offer the potential for near real-time biological tracking and eDNA provides broad, but as yet untapped potential to address eco-evolutionary questions.
Hybrid zones—regions where groups that differ in one or more heritable characters interbreed and produce offspring with admixed genomes—have long been considered important windows into evolutionary processes (Harrison 1990, Harrison and Larson 2014). The study of hybrid zones has provided insights into the nature of species boundaries, the role that hybridization may play in adaptive introgression and speciation, and the influences that climate and environmental disturbance have on the distributions and interactions between species (Harrison 1990, Taylor et al. 2015, Stewart et al. 2016, Taylor and Larson 2019). As the data used to study hybrid zones have shifted towards higher resolution genome-spanning sets of loci (Gompert et al. 2017), we have continued to expand our understanding of the importance and prevalence of hybridization in nature (reviewed in Taylor and Larson, 2019).

Particularly as rates of hybridization increase globally, due to a number of factors which include species introductions, range shifts, and anthropogenic disturbances, many authors have argued for the accurate quantification of hybridization, the examination of temporal trends in extent and location of hybrid zones, and the importance of tracking changes in species interactions at the level of the genome through time (Buggs 2007, Taylor et al. 2015, Grabenstein and Taylor 2018). Although outcomes of hybridization are variable, both positive and negative from an evolutionary or species conservation perspective (Grabenstein and Taylor 2018), without accurate documentation, we cannot determine the consequences of hybridization, or mitigate hybridization in instances where it threatens species survival. We will also be hampered in our ability to understand how anthropogenic change is altering species interactions. Thus, despite renewed calls for temporally repeated and high-resolution studies of hybrid zones, our ability to thoroughly investigate the dynamics within hybrid zones has been limited by various factors.
Most hybrid zone studies are conducted in a single season, across a single geographic replicate. Given our growing awareness that hybridization between the same taxa can have variable outcomes that depend on geography, ecology / life history, local demographics, and habitat, (e.g., Mandeville et al. 2017, Schumer et al. 2017, Stewart et al. 2017), such studies limit our ability to draw broad conclusions about evolutionary and conservation related outcomes of hybridization. While many would prefer to incorporate repeated geographic and temporal sampling into studies of hybridization in nature, the reality of short funding cycles, logistical challenges of geographically replicated field work, and sequencing costs for thousands of samples has limited the number of temporal or geographically replicated investigations of hybrid zones (see Buggs 2007).

Decreased sequencing costs has partially alleviated this problem, even for non-model organisms, bringing such studies within the realm of possibility for most labs. However, geographically replicated sampling at the scale needed to adequately address questions about the consistency of interspecific interactions in hybrid zones remains challenging, especially for organisms that are logistically difficult to directly sample.

**Using environmental DNA to study hybrid zones**

DNA that is collected and extracted from environmental samples is referred to as environmental DNA or ‘eDNA’. It is a means of collecting information without visual observation or direct handling of organisms, which sometimes has negative impacts on the organisms or the habitats in which they live, and requires expertise and spatio-temporal sampling effort (Jerde et al. 2011). Sometimes referring to samples obtained from direct remains (e.g., hair, saliva, scat), much
eDNA work uses indirect genomic remnants found within the environment (e.g., air, water, or soil) which allows for sampling areas of suspected site occupancy and increasing access to habitats that are difficult to sample. Whether subcategorized into intracellular (e.g., DNA enclosed within cell membranes) or extracellular (e.g., free-floating nucleic acids after cell lysis), eDNA represents a biological archive of genes, species, and communities that historically or currently reside within specific habitats. Although challenges remain, a number of studies have successfully (and repeatedly) used eDNA in both aquatic (e.g. Thomsen et al. 2012, Kelly et al. 2014, Pilliod et al. 2014, Deiner et al. 2016, Ma et al. 2016, Stewart et al. 2017) and terrestrial (e.g. Andersen et al. 2012, Ushio et al. 2017, Franklin et al. 2019) habitats for occurrence (presence/absence) and relative abundance measures (number of sequenced eDNA reads) (reviewed in Barnes & Turner 2016, Goldberg et al. 2016, Stewart 2019). Rapid advances in the use of eDNA have also seen non-invasive sampling markers evolve from mtDNA barcodes of various sizes (Foote et al. 2012, Egan et al. 2013, Ma et al. 2016), to diagnostic SNPs (Uchii et al. 2016, Uchii et al. 2017), and nuclear DNA (nDNA; Carpenter et al. 2013, Bylemans et al. 2017, Minamoto et al. 2017, Aylward et al. 2018, Dysthe et al 2018), making the detection of even closely related species, and their potential admixture, possible. Building from recent advances in the use and study of eDNA that expand beyond mitochondrial barcodes, we believe that eDNA is a potentially powerful tool that could augment studies of hybridization and hybrid zones in nature. Studies of hybridization and hybrid zones should use of eDNA to increase temporal sampling (contemporary and historical), to refine and geographically expand sample collection, and to collect data for taxa that are otherwise difficult to directly sample (e.g., rare, cryptic, or otherwise elusive).
Two recent reviews have highlighted new potential uses of eDNA, encouraging a transition from strictly taxonomic monitoring and conservation management, to more ecological (Bálint et al. 2018) and population oriented avenues of research (Adams et al. 2019). We add to this discussion by suggesting that eDNA is a promising but underutilized tool for evolutionary investigations, particularly for studying hybrid zones. The use of eDNA for the detection of macroorganisms was especially significant in monitoring invasive genotypes (Ficetola et al. 2008), which is comparable to documenting parental species genotypes in contact zones. Due to the incredible sensitivity and rapid accumulation of eDNA for occupancy patterns, in near real-time, it should provide an excellent tool for the quantification of low-density, transient, or cryptic species, factors that have traditionally made studying hybrid zones challenging. Ideal hybrid zone sampling frameworks are often difficult to accomplish because many clades along the speciation continuum are poorly understood, including their ecology, phenology, breeding behavior, and how these might differ during divergence; here we argue eDNA sampling may alleviate some of these difficulties.

Expanding the geographic extent and temporal resolution of hybrid zone studies

Collecting DNA from the environment, rather than directly from organisms, can provide high-resolution temporal data across a large taxonomic breadth and geographic context compared to traditional methods which rely on the direct sampling of organisms (Bálint et al. 2018). For rare individuals or cryptic populations (e.g. juvenile forms), low probabilities of detection increase systematic errors and hinder accurate occurrence estimations, but eDNA sampling efforts increase detection rates, reducing false negatives and confirming true absence records (Wilcox et
Further, eDNA collection is both labor, time, and cost-efficient (Qu & Stewart 2019), and the collection of eDNA has frequently been included in citizen science projects (e.g. Biggs et al. 2015, Buxton et al. 2018), or accomplished via extensive collaborative networks (Wilcox et al. 2018). These aspects alone would vastly improve both the geographical extent and temporal resolution of sampling across hybrid zones, particularly for complex mosaic hybrid zones (e.g., Larson et al 2013) or hybrid zones that extend across national borders (e.g. Stewart et al. 2016, Ryan et al. 2018) and/or have broad geographic distributions (e.g., Scriber 2011). The ease of collecting environmental samples (e.g. water or soil) further means that dense geographic and repeated temporal sampling could refine known hybrid zone boundaries and identify new regions of contact, while simultaneously allowing for broader sampling coverage without being prohibitively expensive or labor-intensive.

Moreover, although eDNA often degrades rapidly in nature (on the scales of days to weeks) making it an ideal tool to monitor the contemporary distribution of organisms (Goldberg et al. 2016), eDNA can be successfully amplified up to 1 million years after it is shed into the environment (Willerslev et al. 2007, Kirkpatrick et al. 2016). When combined with dating methods (e.g., isotopic analysis, rare historical events that leave paleoecological traces, or annual lamina in sediments; reviewed in Bálint et al. 2018), eDNA may illuminate the historical spatial legacy from species movements. For example, a recent study successfully used eDNA to identify a historical invasion front, contrasting the ecological impact of the invasive species to recent climate change events (Ficetola et al. 2018). Importantly, even the contemporary collection of eDNA can allow for a retroactive look at spatial patterns of occurrence and relative abundance in genes and species through time, which has obvious application to the study of hybrid zones.

Aspects of hybridization history and hybrid zone movement, which are often difficult to deduce
(e.g., the source and speed of genetic invasion fronts [i.e. hybrid zone movement], the frequency of reticulated contact, or establishment of tension zones), could all be addressed using spatially and temporally explicit eDNA collections.

Making predictions about hybrid zone movement may be possible when using eDNA tools for hybrid zone investigations. Species distribution models (SDMs) can link biological observations, geospatial habitat, and climactic covariates to forecast future distribution probabilities based on eDNA data (Muha et al. 2017, Wilcox et al. 2018). By using similar techniques, one could geographically sample hybrid zones, along with the abiotic and biotic parameters that they are correlated with at high-resolution, and then predict ecologically realistic patterns of introgression and movement trajectories through time. This is an especially useful opportunity for analysing dispersal pathways (Muha et al. 2107) as introgression from introduced species (e.g., Hohenlohe et al. 2013) and climate change (see Taylor et al. 2015) alter species interactions and distributions.

Providing insight into cryptic aspects of hybridization and ecology

We further think that eDNA can serve as a springboard for the collection of otherwise difficult to sample data. Although our current understanding is that eDNA derives from both dead (e.g., Dell’Anno and Danovaro 2005, Pietramellara et al. 2009) and living (e.g., Pochon et al. 2017) biomass, quantifying viability and fecundity dynamics within hybrid zones might also possible with eDNA. Sources of genetic material within environments are varied (Stewart 2019) and intracellular or eRNA sources are assumed to originate from metabolically-active living organisms before being rapidly removed from the environment. Examination of proportions of
eDNA / eRNA (e.g., Steven et al. 2017), or intra to extracellular (correcting for degradation),
could allow for inferences related to general patterns of mortality either due to hybridization
itself, or via species interaction and competition. If sampling is directed at discrete life-stages
(e.g. egg, larval, and adult forms) that occupy distinct temporal (e.g. seasonal) or geographical
(e.g. terrestrial vs. aquatic, or species-specific aquatic vertical distribution of gametes or eggs;
Stewart 2019) realms, eDNA collections may also open windows into differential mortality
throughout development.

The detection of eDNA is also known to spike in aquatic environments during
reproductive seasons (e.g., Laramie et al. 2015, Spear et al. 2015), with breeding events
characterized by higher nDNA relative to mtDNA, facilitating quantification of reproductive
bouts within hybrid zones and 2) phenological breeding patterns in the parental species coming
into contact. As X- and Y- linked markers (e.g., Taberlet et al. 1993; Brinkman & Hundertmark
2009), and sex-associative mtDNA heteroplasmy markers (Mioduchowska et al. 2016) have also
been developed for non-invasive sampling, it may also be possible to determine sex ratios within
populations that have genetically determined sex. This is especially important for species that do
not display sexual dimorphism. Sex-linked markers could further provide insight on postzygotic
reproductive isolation, such as hybrid dysfunction (Haldane’s rule). Likewise, eDNA would
allow the quick retrieval of diagnostic genes that differ between the parental species within
hybrid zones when accompanied with high-quality reference genomes and initial exploratory
work.

Practical considerations and potential limitations
When coupled with proper marker design and sampling strategy, eDNA is robust with low error rates, yet it is not without its limitations. For example, most eDNA studies to date have employed mtDNA as their marker of choice, allowing for the delineation of maternal lineage or contact boundaries, but failing to incorporate aspects of admixture. This is not a hindrance for exploratory analyses, as a necessary first-step would be delineating species boundaries and identifying potential geographic areas of introgression (Fig. 1). The relative proportions of genetically similar taxa can be quantified using mtDNA SNP detection via eDNA sampling (e.g., Uchii et al. 2016, Uchii et al. 2017), but key information regarding the dynamics of species interactions, such as hybridization, would remain unavailable. However, eDNA collections quantifying nDNA have now been used successfully in the field (Minamoto et al. 2017, Dysthe et al. 2018), and could reveal important spatio-temporal patterns in areas of contact. By combining different markers, researchers could perform population level cline analysis, comparing expected proportions of inheritance for each marker (Fig. 1). As the pool of eDNA data would potentially represent an amalgamation of all individuals sequences within a population, analysis would be akin to Pool-Seq pipelines (e.g., Pfenninger et al. 2105, Taus et al. 2017). If diagnostic markers show strong evidence for cytonuclear discordance via eDNA surveys, subsequent individual assessments should be made using traditional sampling. Although individual measures of hybridization (F1, F2, and backcrosses) are, at present, beyond the scope of eDNA, exploratory analysis using eDNA would decrease guess-work in geographic sampling, potentially helping to pinpoint populations of importance. For refining and expanding sampling for well-studied hybrid zones with a priori information about admixture, eDNA represents a potentially valuable addition to current sampling protocols. We emphasize that eDNA, like many other tools, should not be used as a standalone method for the study of hybridization and
hybrid zones in nature. eDNA has always been a compliment to other traditional sampling practices, whether for biodiversity monitoring to confirm presence/absence assays, or in this case, to clarify levels of admixture.

*Environmental DNA is a currently underutilized tool for studying hybrid zones*

Genomic data in the environment offer the potential for near real-time biological tracking. Since its inception for macroorganismal use (Ficetola et al. 2008), eDNA has been widely adopted and utilized in conservation biology, although it provides broader yet untapped potential to address eco-evolutionary questions. eDNA is especially useful for detecting cryptic species and unique genotypes. Thus, a promising application for eDNA in an eco-evolutionary framework is to obtain quantitative measures of species presence / absence and to link this to the chronology of spatial occurrence and relative abundance. eDNA could facilitate the reconstruction the historical presence and movement of species boundaries (and hybrid zones) with future research avenues including investigating species boundaries, delineating fine-scale hybrid zones, and tracking the spatio-temporal introgression of invasive genotypes. Importantly, eDNA allows for the data and original sample to be stored within a repository, archived so that new questions may be asked or new taxa may be studied. The significance of this cannot be understated given the rapid discovery of new markers or genes under selection, rendering eDNA an invaluable tool for evolutionary studies, now and in the future.
Acknowledgements

The authors would like to thank Erica L. Larson, Catherine E. Wagner, Stephen C. Lougheed, and Jun Ying Lim for invaluable input on early versions of the manuscript.

References


Franklin TW., McKelvey KS., Golding JD., Mason DH., Dysthe JC., Pilgrim KL., et al. 2019. Using environmental DNA methods to improve winter surveys for rare carnivores: DNA from snow and improved noninvasive techniques, Biological Conservation, 229: 50-58


**Figure 1:** Examples of how spatial and temporal eDNA sampling could facilitate hybrid zone research, including expanded geographic replicates, population-level cline analysis (mitochondrial DNA, mtDNA; nuclear DNA, nDNA), and comparisons of contemporary and historical samples for the detection of unknown species distributions. Diagram key is located in the top left corner.