k-mer analysis shows hybrid hummingbirds perform variable, transgressive courtship sequences

Brian M. Myers[1] , David T. Rankin[2], Kevin J. Burns[1], Alan Brelsford[2], and Christopher J. Clark[2*]

[1] Department of Biology, San Diego State University, 5500 Campanile Drive, San Diego, California, 92182;

[2] Department of Evolution, Ecology, and Organismal Biology, Speith Hall, University of California, Riverside, California, 92521 USA

*Corresponding author: cclark@ucr.edu

Acknowledgements: We are grateful to Elisa Henderson and Nadje Najar for feedback on the manuscript. Brandon Lebeiko, Aubtin Rouhbakhsh, Michael Jimenez, Zachary Williams, Lisa Tell, and CJ Ralph, for assistance in the field and with analyses. We thank Gwen Baluss, David Payne, Sam Cuenca, Bree Yednock and Angela Doroff with the South Slough National Estuarine Research Reserve. For research permits, we thank Sara Griffith and Oregon State Parks; Nita Barve, Amber Transou, Carol Wilson and the California State Parks, The USGS Bird Banding Laboratory; the California Department of Fish and Wildlife, the Oregon Department of Fish and Wildlife, Kip Wright and the Bureau of Land Management staff at New River, the Alaska Department of Fish and Game, and the United States Fish and Wildlife Service.

Conflict of interest: The authors of this manuscript certify that they have no conflicts of interest.

Author contributions: BMM: Methodology, Formal Analysis, Investigation, Data curation, Writing, Original draft, Visualization, Funding Acquisition. **DTR**: Formal Analysis, Investigation. **KJB**: Investigation, Supervision, Funding acquisition. **AB**: Investigation, Supervision, Funding acquisition. **CJC**: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft; Writing - Review & Editing, Visualization, Supervision, Funding Acquisition.

Funding statement: Funding was provided by the National Science Foundation (IOS-1656867 and IOS-1656708). Additionally, funding was provided by Society of Integrative and Comparative Biology, the American Ornithological Society, the Los Angeles Audubon Society, the Pasadena Audubon Society, the American Museum of Natural History, the California Doctoral Incentive Program, the Crouch Scholarship for Avian Behavior, and the James and Mary Crouch Memorial Scholarship.

ABSTRACT [281 words]

Behaviors are challenging to describe. Here, we apply k-mer analysis to characterize complex courtship behaviors at four hierarchical levels: elements, displays, bouts, and repertoires, on two species of hummingbirds and their hybrids. During courtship, male Rufous Hummingbird (*Selasphorus rufus*) performs three types of displays: shuttles (S), half pendulums (H) and rufous dives (R), while Allen's Hummingbird (*S. sasin*) performs four displays: S, H, Allen's dives (A), and the pendulum (P). Here we explore the behavioral sequences expressed by their hybrids, on a sample of displays from 35 Allen's and 46 Rufous sampled far from their hybrid zone, compared against 306 wild male birds from a hybrid zone in Oregon and California. Among Allen's, hybrids, and Rufous, there was almost no variation in the kinematic elements that comprised displays. By contrast, the displays themselves, display sequences, and repertoires varied substantially among these three groups. Some hybrids performed transgressive display variants in which kinematic elements of typical displays were missing or duplicated. This transgressive segregation implies the displays are under oligogenic control. We analyzed display sequences using k-mer analysis, in which we binned display sequences into 3-mers (sequences of three). Typical 3-mers for Allen's included sequences PPP, PPA, and SHP while typical 3-mers for Rufous included SHR and RRR. The most defining characteristic of certain hybrids was their sequence variability: the display repertoire of some hybrids encompassed the repertoire of both parent species (e.g. performed both PPA and RRR), as well as 3-mer sequences never observed in either parent such as PRA. Such high within-individual variability in hybrid phenotype differs from how modular morphological traits (such as flowers, feathers, or hair) are expressed in hybrids. This within-individual variability may be unique to behavior.

Highlights:

-We analyzed courtship display sequences from 306 hummingbird hybrids within a hybrid zone

-Some hybrids performed transgressive displays that included extra or missing display elements

-Some hybrid individuals were highly variable in their courtship display sequences

Keywords: admixture, courtship, ethology, modularity, k-mer, transgressive segregation

INTRODUCTION

One of the early goals of comparative ethologists was to make ethology similar to the study of comparative morphology, such as by making homology statements that would permit comparisons of how behavior evolved in related taxa (Lorenz 1941). Postulating homology requires identifying natural units of the phenotype (Baerends 1958; Barlow 1977), which, for behavior, can be challenging (Wagner 2014; Wenzel 1992). This search for natural units led to the development of the concept of the "Fixed Action Pattern" (FAP), a relatively invariant sequence of motor acts that, once initiated, tends to run through to completion (Thorpe 1951). This concept emphasized invariance and independence of a relatively complex series of motor acts (Lorenz 1950, 1958; Schleidt 1974): the FAP as the granular ‘molecule' of behavior. But most behaviors are difficult to pin down in this way: even ‘good’ FAPs are rarely completely invariant or completely free from feedback (outside inputs that affect expression).

The FAP is not the only phenotypic level at which behaviors can be studied. Behavior, like morphology, is hierarchical in its structure. Just as molecules are made of atoms, complex behaviors such as FAPs can be broken into smaller individual, interdependent parts (Baerends 1958; Barlow 1977; Barlow 1968), which here we call *kinematic elements*. As an analogy, if a display (FAP) is a word, the kinematic elements are the letters of that word. And if a display is a word, then at a higher level, sequences of displays that are strung together into bouts are akin to sentences. Finally, bouts are also performed repeatedly (Elias et al. 2012; Patricelli, Krakauer & Taff 2016; Scholes 2008) and often differ from one to the next; the sum total of an animal’s phenotype (including both the composition and sequencing of displays), across all bouts, comprises its display repertoire and sequence repertoire. Thus the four levels of behavior we consider here, from smallest to largest, are: *elements*, *displays*, *bouts*, and the *repertoire*.

Here, we examine how elements, displays and bouts are put together, and vary, to form the repertoire of courtship displays of male hummingbirds from a recently described a hybrid zone between Allen’s and Rufous Hummingbirds (*Selasphorus sasin* and *S. rufus*) (Myers et al. 2019). One noteworthy difference in their courtship displays is Allen’s hummingbird performs a display called the ‘pendulum display’ (figure 4.3 in Clark 2016), which we symbolize here as P (described below), and Rufous does not. Both species produce a close-range shuttle display (symbolized S) in which the male flies back and forth right in front of a female in a bouncy flight, making pulses of wing sound; and a larger dive in which the male ascends and then descends at high speed while making sounds with his tail (Clark 2009; Clark et al. 2018).

The form of the Allen's dive (A) and Rufous dive (R) is similar. One difference is the two species make different sounds with their tail-feathers (Myers et al. 2019), which we hereafter ignore. The dives differ in one kinematic element, 'writhing', in which the male makes sounds with his wings while flipping his tail up and down. This element is present at the start of an Allen’s dive as he ascends to dive, but absent from this part of Rufous dives; both species also do writhing at the end of the dive. The form of the shuttle seems to be identical between the two species (Clark 2014; Myers et al. 2019). Finally, Myers et al. (2019) described another display, the “half pendulum” (H) which both species perform after shuttle displays, i.e. this display is nearly always performed as (SH), although, very rarely, H can be performed in isolation (see results). Thus, Allen's has a *display repertoire* of (S, H, P, A) and Rufous has (S, H, R), where A and R are nearly the same, differing only in a single element.

Assuming expression of these displays is predominantly under genetic control, if the displays are oligogenic (controlled by a small number of genes), then we predicted that we would find occasional transgressive behaviors manifested in some hybrids. Transgressive traits are characters that fall outside the range of variation of expression in either parent, as the result of interactions between alleles inherited from each parent (Lexer et al. 2003). Another hypothesis is hybrids may express atavistic behaviors, which are behaviors not present in either parent species but present in an outgroup. Lorenz (1958) reported that crosses between Chiloe Teal and Bahama Pintail performed atavistic display sequences.

A prime way that courtship display behaviors vary is in their sequence (Elias et al. 2012; Kershenbaum et al. 2016; Scholes 2006, 2008). "ON", "NO", and "NOON" are different words due to sequence differences, not the identity of their component letters. Similarly, sequences of displays within a bout differ between Allen's and Rufous. Myers et al. (2019) found that Allen's performs an average of 10 pendulums before any dive, i.e. PPPPPPPPPPA is a typical Allen's bout, and Allen's rarely performs more than one dive in sequence. Rufous Hummingbird perform an average of 4 consecutive dives in a bout, such that RRRR is a typical Rufous sequence. To describe this sequence variation, we considered Markov models (Kershenbaum et al. 2016), but these were not best-suited to our dataset because the hummingbird displays violated the Markovian assumption of 'memorylessness'. Allen's Hummingbird displays appeared to exhibit 'memory': the probability of transitioning from P to P declined and P to A increased over the course of a bout; moreover, accurately estimating transition probabilities of rare transitions can be error prone. Instead, we adopted a related sequence analysis method that places less emphasis on transition probabilities, and is already widely used for analysis of DNA sequences: *k*-mer analysis. To our knowledge, our study is the first to apply this method to study behavioral sequences.

A *k*-mer is a sequence that is *k* components long. For instance, a 3-mer is a sequence of 3, such as RRR or PPA (i.e. a 3-mer is akin to a 3rd order Markov model). We divided displays of Allen's, Rufous, and their hybrids into kinematic elements (symbolized with lower case italic letters, *a*, *b*, *c* etc.). Then using these elements, we constructed an 'alphabet' of displays, symbolized with six primary upper-case letters (A, E, H, P, R, and S). When H, P, R, and A displays arose that deviated from our base definition, we assigned every variant of a display a unique subscript (such as A_A, A_L or A_X) within this alphabet. We then implemented *k*-mer analysis to describe the differences in sequences of this 'alphabet' for the display sequence repertoire. On a sample of 386 wild birds, comprising male Allen's (N = 35), Rufous (N = 46) and individuals from the hybrid zone (N = 306), we addressed the following questions: how are hybrid display sequences structured relative to the parents, and how do hybrids vary from each other and from parental phenotypes?

METHODS

Ethics statement

This research was conducted in compliance with the ASAB/ABS Guidelines for the Use of Animals in Research, the IACUC at the University of California, Riverside (protocols 20130018, 20160039, and 20190019), USGS Bird Banding Permit #23516, and the following collecting permits: USFWS permit #MB087454-1, California Department of Fish and Wildlife permit #SC006598, California State Parks permit #17-820-01, Oregon Department of Fish and Wildlife permit #096-19, #074-18, #055-17, #049-16, #082-15, and #103-14, Oregon Parks and Recreation Department permit #011-14 and Alaska Dept. of Fish and Wildlife permit #18-126.

Most individual males sampled in this study were banded to identify them individually and avoid re-sampling them between years. A small number of males were deposited as specimens in the San Diego State University Museum of Biodiversity (SDSU) and the San Diego Natural History Museum (SDNHM) as a part of another study (Myers et al. 2019). As male hummingbirds are not known to directly participate in nesting, removing males should have minimal direct impacts on nearby nesting females. We elicited displays by placing a wild-caught, caged female on male territories. We sought to minimize adverse impacts on populations by minimizing the number of females held captive (1 to 5 per year). This meant that, in some cases, the same female was used to elicit displays from many males. Females were monitored for stress while in captivity, and were fed ad lib with a complete hummingbird diet (Nektar Plus, Nekton GmbH, Germany). Females were usually released at the point of capture at the end of each field season, or a few were humanely euthanized (as a part of another study).

Sampling

We sampled courtship data from male Allen’s and Rufous in allopatry; and their hybrids within and outside of the hybrid zone from March-June of 2014-2019 and 2021. Our sample included 183 birds described in Myers et al. (2019), and an additional 203 birds. The 46 birds we *a priori* labeled as allopatric Rufous were sampled far from the hybrid zone, including from three inland populations (> 100 km from the coast) in Douglas County and Jackson counties, Oregon, a coastal population in Clatsop, Oregon; and two populations in the vicinity of Ketchikan and Juneau, Alaska (see fig. 2 in Myers et al 2019 for a range map of the two species, but note that certain populations, e.g. in Alaska, are not shown). The N = 35 birds we defined as allopatric Allen’s Hummingbirds were mostly sampled far from the hybrid zone, including 27 Migratory Allen's (*S. sasin sasin*) from along the coast in California and eight

Sedentary Allen's Hummingbird (*S. sasin sedentarius*) from southern California. The furthest north population was n = 10 individuals from Humboldt Redwoods State Park in southern Humboldt county. Analyses in Myers et al. (2019) suggested that no individuals in this population had hybrid phenotypic characters (Fig. 2 in Myers et al. 2019). The N = 306 birds sampled between Lincoln County, OR and Humboldt County, CA (north of Humboldt Redwoods State Park) were *a priori* labeled hybrids, including all birds sampled inland along the Klamath River on the California side of the CA-OR border (in Siskyou county, CA).

This definition of ‘hybrid’ is necessarily loose. Many of these birds labeled as 'hybrid' were sampled from the periphery of the hybrid zone. Although these populations appeared to contain some individuals expressing one or more hybrid phenotypic characters, populations at the periphery generally contained individuals that were phenotypically indistinguishable from Allen’s or Rufous. Although we have DNA for all of the birds described here (Myers et al. 2021; authors in prep), due to extensive introgression and recent divergence, few genetic markers distinguish between these species (Henderson & Brelsford 2020). Therefore, the conservative approach is to consider all birds from within the geographic regions described above as ‘hybrid’, even though some were behaviorally and morphologically indistinguishable from a parental species, since diluting the pool of ‘true’ hybrids with these birds is conservative relative to the conclusions presented.

We sampled male behavior by taking audio recordings of courtship displays. Audio recordings of wild hummingbird displays are far easier to obtain than video, and more reliable than visual observation. Sound recordings permit quantification of male display behavior sequences because males produce a specialized sound with modified outer wing-feathers when they fly, termed the wing trill, in which one pulse of sound is produced by each flap of the wings (Clark 2016; Clark & Mistick 2018). Other qualities of the trill (e.g. loudness) provide additional behavioral context, such as whether the bird is ascending for a dive.

Most individual behaviors on which we focused were diagnosable from a sound recording alone. A few rare display variants were not individually diagnosable from sound, because the recording alone sounded identical to a more common behavior. To account for these rare behaviors we spoke to the microphone in real time to annotate these behaviors when they were visible. However, as we frequently lost visual track of the birds as they displayed (but could still hear them), it is virtually certain that we have under-counted the presence of certain rare behaviors, since these were scored as the more common display variant.

One behavioral element, ‘writhing’, was present in A (twice), H, P and R displays. In writhing, a bird produced discrete pulses of the wing trill, termed ‘chirrups’ (Myers et al. 2019), which we quantified by counting the number of sound pulses on a spectrogram in the program Raven Pro 1.5 (Charif, Waack

& Strickman 2008), when this feature was visible. For the remaining elements we considered, we assessed their presence/absence from spectrogram. In some instances we visually observed a bout of display and recorded its sequence in our written field notes, but were not able to obtain a sound recording. All birds had sound recordings of some of their bouts of display.

Among the bouts we recorded, the focal male had to perform at least one shuttle (S) and one dive (A or R). We usually attempted to record at least 10 display bouts per bird, except for pilot birds recorded in 2014, prior to establishment of that criterion. If after 10 bouts the bird had not performed a shuttle or a dive, we continued to sample additional bouts until they were observed. (All birds eventually did perform the missing display type). We defined a bout of display as a sequence of displays that concluded when a male ceased to display for at least two seconds (usually, the break between bouts was much longer). Further, a display bout had to last at least two seconds to be counted (displays shorter than this were shuttle displays, which did not vary geographically in the birds we sampled).

Captive females elicit displays

Most bouts of display were elicited with a female *Selasphorus* hummingbird in a mesh cage, nestled on top of the thick bushes typical of a male territory, in a location visible to the recordist. There were three benefits to using a caged female. First, she provided a somewhat standardized stimulus that couldn't leave mid-display. "Natural" females have presumably interacted with the territorial male previously. We hypothesize males modulate their displays in response to the nesting status of females. Females likely nested somewhere in the thick bushes that characterized typical male territories, but as females were very hard to find/observe, when recording a male, we almost never knew whether there was a female nearby, let alone her nesting status. Moreover, male displays to wild females were frequently truncated, terminated prematurely when the wild female departed in the middle of a bout. In Allen's hummingbird, this tends to reduce the number of dives recorded, since dives are typically at the end of a bout. Although bouts may be truncated for other reasons we could not control (especially, another male hummingbird interrupts the focal male), use of a caged female eliminates this one source of variation. The second advantage of a caged female: males that did display often performed several bouts of display in a short period of time, facilitating acquisition of a large sample size. Third, displays were performed close and visible to the recordist. By contrast, natural displays were often distant and not visible (e.g. behind a bush) to the recordist. As male territories were often densely packed, it was often hard to tell whether it was the focal male that performed distant, unseen displays. In sum, "natural" displays likely had several extrinsic sources of variation, which use of a caged female to elicit displays controlled for.

The downside of using a caged female to elicit displays is that nearly all males became acclimated (i.e., ceased displaying to her), sometimes after only performing a few bouts of display. Unlike a wild female, a caged female did not arrive by flying across the male's territory. Once spotted by the male, a caged female could neither flee nor attack the male (two ways females seem to signal disinterest). While some males readily performed >10 bouts of display, we suspect that males generally acclimated because they decide to display or not in response to subtle signals given by the female, or in response to natural context (e.g. a female attempts to feed on a male's territory). A caged female, involuntarily present on the male's territory, lacks this context, so males would cease to display. Moreover, as we sought to minimize the number of females affected by this research, we put the same female on different male territories. Female responses to males change with time: females that had been in captivity for a while learned to remain motionless when a male was around, preventing him from detecting her.

We were not able to systematically score to which stimulus each display was directed, because the recordist's attention was typically focused on the male, not the recipient of displays. It was common for there to be multiple possible recipients: wild females often surreptitiously approached the caged female (perhaps to expel an intruder from the territory), initially un-noticed by us. Whereupon the resident male began to display in the vicinity of the caged female, there were many times we were unsure whether the display was directed towards a wild female or the caged female. Moreover, since we captured males (for another study) by placing a hummingbird feeder their territory, sometimes the male displayed to other birds, often of unknown sex, that approached the feeder.

A goal of ten bouts was chosen as a compromise between the need to sample each male's display repertoire as fully as possible, and the practicality that it was fairly typical for a male to cease responding after 7-10 displays. As a male lost interest in a caged female, he generally either stopped performing bouts of display to her entirely, or switched to performing just short shuttle displays. Some males did not respond to a caged female at all. As recording natural displays from these males was tedious, few of these males were sampled. For instance, to address repertoire size in Allen's, in spring 2021 we recorded >60 bouts performed to natural stimuli from two Allen's Hummingbirds in a yard in Riverside, CA. Each of these males required several weeks to record (neither would display to a captive female).

We used a Sennheiser (Wedemark, Germany) MKH 70 shotgun microphone or MKH 20 microphone in a Telinga Pro parabola, an Audio-Technica AT875R shotgun microphone (Stow, Ohio), a Tascam DR-05 portable recorder (Tokyo, Japan, sampling rate 44.1 kHz, 16 or 24 bit), with a Tascam DR-60DmkII audio recorder (Tokyo, Japan, sampling rate 44.1 kHz, 16 or 24 bit), and a Sound Devices 702 24-bit digital recorder (Reedsburg, Wisconsin, sampling rates: 44.1, 48, or 96 kHz). Spectrograms were analyzed with FFT windows of 1881, 2048, or 4096-sample to account for the respective sampling rate differences. All analyses of behavioral elements, displays, bouts of display, and repertoire were performed in R Studio v3.4.3 (RStudio Core Team 2019), R v3.5.0 (R Core Team 2018), while other statistics were performed in JMP 15.0.

Display nomenclature

Elements: the building blocks of displays

The elements that comprise the displays of Allen's, Rufous, and their hybrids are based on those described in Myers et al. (2019): *a*) short descent, *b*) short ascent, *c*) writhing, *d*) long ascent, and *e*) long descent (Figure 1, Table 1). It was necessary to make three small changes to the scheme presented in Myers et al. (2019). First, describing certain rare display types not encountered by Myers et al. (2019) required defining an additional behavioral element to distinguish between them. Specifically, we added "turns around" (behavioral element *g*) to account for when a male changed direction between one display and the next, since we encountered a small number of hybrids that occasionally omitted this element. Second, we also defined a new version of the element 'ascent' to account for the mini-dive display type, described below. All other display variants (below) could be described using the kinematic elements named in Myers et al. (2019).

Third, the shuttle display was a nuisance variable: it was integral to male repertoires, meaning it was essential to include, but was not informative with respect to the goals of the study, since all birds (Allen's, Rufous, hybrids) seemed to perform identical versions it, and transitions to/from the shuttle did not vary between species. The number of individual back-and-forth repeats of the shuttle (element *f* in figure 1) was highly variable (but did not seem to differ among the species), so instead of counting these repeats, we treated every shuttle display as a single, discrete display (S), irrespective of the number of repeats.

Displays and display variants: the alphabet

Our definition of displays (FAPs) build on the displays types presented in Myers et al. (2019). Symbolized with upper-case letters, the main types of displays were: S, P, H, A, and R. For reasons described above, there were no variants of S. P, H, R and A had a standard sequence of elements that comprised the 'normal' version of the display (Table 2). When we recorded a display that departed from the standard sequence of elements, we gave that variant a subscript letter, e.g. A_X, A_A, A_L. From the sound recordings, bouts were scored by assigning elements to each sound within the recording, then assembling these element sequences into the display types (the 'letters' of the alphabet), and from this, the sequences of displays that comprised the bouts. We performed a principal components analysis on the frequency of each display variant for each bird, to describe variation in the prevalence of display variant.

Display sequences & sequence analysis

To analyze display sequences, we only included the alphabet of the main five letters plus the dummy letter E, so that the full alphabet for sequence analysis comprised (A, E, H, P, R, S). We subsumed each display variant into the display type that it seemed to be derived from. The dummy letter E (for 'End') was placed twice at the beginning of the first bout, and twice at the end of each bout. We then concatenated each bird's bouts into a single sequence, the sampled repertoire. For example, for a bird with two bouts, first a shuttle - half pendulum (SH) and then two Rufous dives (RR), the concatenated sequence was "EESHEERREE". Since we used k = 3 (described next), adding E twice in between each bout has the effect of making each bout independent from the others in the analysis.

We applied a k-mer length of k = 3 to the concatenated sequences for each bird. This means that we split each bird's repertoire into sequences of 3 (EESHEERREE has the following 3-mers: EES, ESH, SHE, HEE, EER, ERR, RRE and REE). One reason for k = 3 was practical: since the possible number of k-mers given an alphabet of N is N^k, k = 3 was more feasible for the dataset than a higher order number. The other reason was empirical: while k = 2 is a nearly trivial sequence length, k = 3 was sufficient to capture much of the sequence variation that we observed to differ between Allen's and Rufous.

Each male's sampled repertoire was a concatenated sequence of letters. From these, we calculated a k-count matrix, which is a matrix of counts of all 3-mers within each repertoire across all the birds, using the *kmer* package (Wilkinson 2018) in R v4.1.0. For instance, considering the sequence of 'PPPPPPPPPPA' has an alphabet of 2 (P and A), has a k-count matrix of (PPP: 8, PPA: 1, PAP: 0, APP: 0, PAA: 0, APA: 0, AAP: 0, AAA: 0), i.e. out of the 2^3 possible 3-mers, 89% of the present 3-mers are PPP, 11% are PPA, and the rest have a frequency of 0%.

Given our alphabet of 6, there were theoretically $6^3 = 216$ 3-mers. Given our definitions, 32 of these were not possible. For instance, SS* and *SS 3-mers were not possible due to how we defined S; and *E* (for * =! E) was not possible since E by definition always occurred in pairs. We removed from the matrix all hypothetical 3-mers that were not present in any individual, yielding a total of 129 3-mers that were present in at least one individual. The k-mer frequency within the k-count matrix had a skewed distribution: two 3-mers (PPP and RRR) had comparatively high counts within certain individuals due to the way that P and R displays were repeated, while virtually all remaining 3-mers had within-individual counts of < 10, and most 3-mers had counts of 0 in most individuals. Since the number of bouts varied widely among birds, we normalized by dividing the 3-mers counts by the number of bouts performed by the bird (after excluding bouts of only shuttle displays). We then calculated Principal Components on this bout-normalized k-count matrix.

Estimated true repertoire

Each animal's 'true' sequence repertoire, the 3-mer sequences it would perform in an infinite number of bouts of display, is unknown. To assess sampling and sequence repertoire, we constructed 3-mer accumulation curves of the cumulative total number of 3-mers each individual expressed as a function of measured numbers of bouts, with the *specaccum* function within *vegan* R package (Oksanen et al. 2020). We also calculated the number of singleton 3-mers, which are 3-mers that a given individual bird only expressed once within its sampled repertoire.

Data availability

A spreadsheet of all of the display data presented here is included in the supplemental materials (Supplemental table S1). A small sample of sound recordings has been uploaded to the Macaulay Library (Supplemental table S2), and the entire set of recordings [will be uploaded to Dryad upon acceptance of this paper].

RESULTS

We recorded a total of 4,399 bouts of display from N = 387 wild male hummingbirds: 35 allopatric Allen's, 46 allopatric Rufous, and 306 birds from the hybrid zone. Of the four levels of hierarchy considered (elements, displays, bouts, and sequence repertoire), there was relatively little variation in kinematic elements themselves, but there was substantial variation in displays birds performed due to insertion, deletion, and rearrangement of these *kinematic elements*. We detected several rare display variants performed by hybrids. In most instances, these rare display variants were displays in which individual kinematic elements were inserted or deleted relative to the typical form of the display. The variation included transgressive displays; substantial variation in display sequences, including transgressive sequences; and substantial variation in repertoire size: many hybrids had more variable displays than either parent species.

Kinematic elements varied little

Of the four levels we considered, kinematic elements showed the least variation. We detected rare variants of displays (described further below), but only two rare variants required us to define kinematic elements that differed from the scheme we used in Myers et al. (2019). One type of rare variants we accommodated by defining a new ‘turns around’ element (*g*), as described above. Another rare display entailed naming a new element, *d’*, “intermediate ascent”, defined as an ascent of about 10 m, intermediate between *b* (short ascent of ~2m) and *d* (long ascent of ~20 m). This element appeared within the ‘mini-dive’ display (see below). It is possible that in birds expressing the mini-dive, height ascended varied continuously from 2 m (the height of a typical element b) all the way to > 20 m (the height of element d). However, estimating dive height quantitatively in the field is difficult and so we did not have quantitative estimates for most birds in this study. Therefore, we binned this variable into 3 states: *b* (~2 m), *d’* (~10 m) and *d* (~20 m).

One result was the absence of an effect: with the rare exception of the 'intermediate ascent' just described, we did not detect any other differences in the composition of elements themselves between either parent species or any hybrid. Moreover, all 387 individuals expressed all of the ‘regular’ display elements at least some of the time; no elements were entirely or consistently missing in any individuals.

For example, of the small number of individuals that expressed the 'mini-dive' phenotype (ascending for dives only half the height of a normal dive) all also expressed normal dives; the 'mini-dive' did not replace normal dives. Or, hypothetically a hybrid could omit the ‘writhing’ element from all displays that normally contained it (dives, half-pendulums and pendulums). We did not find any examples like this. While certain individual hybrids sometimes omitted writhing in hybrid display variants we observed (see below), that same bird inevitably performed normal writhing in another iteration of that display. Thus it was not the element ‘writhing’ itself that was missing in that individual. Rather, the variation was at higher levels of integration within the phenotype of the individual.

Displays

Summary: Collectively the birds performed a total of 20,901 displays. There were five main categories of displays: shuttles (S), half-pendulums (H), pendulums (P) Allen's Dives (A) and Rufous Dives (R) (Table 2, Figure 3) which we describe in more detail below. We did not consider variation in the shuttle. Among the dives, half pendulums and pendulum displays, we recorded 18 variants. Nearly all of these comprised insertions or deletions of elements, relative to the 'typical' form of the display. In Table 2 and Fig. 3, insertions are <u>underlined</u> and deletions are indicated with an underscore, _. Only the "mini-dive” (A_M and R_M) displays required a modified element. Many of these display variants were rare and unique to birds within the hybrid zone (Table 2).

Shuttles (S). All birds performed shuttle displays. Although this display is integral to courtship, e.g. males shuttle to females right before attempting copulation (Clark & Mitchell 2013), this display also seemed to be used in aggressive or agonistic interactions performed to intruders on a male’s territory. In addition to females, males directed shuttle displays towards passerine birds (e.g. bushtits or warblers), squirrels, male conspecifics, heterospecific hummingbirds, and inanimate objects (such as a hummingbird-sized leaf). Shuttle duration was variable and seemed tied to male motivation. As males became acclimated to the female in a cage, it was somewhat common for the male to cease performing other display types and to only perform shuttle displays of short duration. Also, at the end of the breeding season, some males ceased performing other displays and only performed shuttles. For reasons explained above, we ignored the variation in duration of shuttle displays, thus there was only one type of shuttle.

Half-pendulums (H). Half-pendulums were nearly always performed right after a shuttle display. This display was frequent in Allen’s, Rufous, and hybrids. This display was identical to the element sequence of the final part of a typical dive, or the second half of the pendulum display (hence, ‘half’ pendulum). We defined this as a separate display from the shuttle because many shuttle displays did not end with a half-pendulum. While a few individual birds did not perform this display within the sample of displays we recorded, we infer that this was likely due to limited sampling, rather than true absence. Rarely, this display was not performed at the end of a shuttle display. The context in which H was not performed after a shuttle were instances in which the bird flew up to the female as if to shuttle, but then skipped the shuttle and went straight to a half-pendulum (n = 4 birds); or instances in which a shuttle display was interrupted, there was a pause of >2 seconds (for instance, a male shuttled, then fed from a hummingbird feeder), then resumed displaying with H (n = 3 birds). As a pause of 2 seconds was our operational definition of a new bout, the bird began a bout with H.

There were three variants of H, which we symbolized as H_D, H_X, and H_W (Table 2). In H_D, the male performed an H that was not followed by a turn-around, but rather, continued to ascend for another behavior such as a dive (type A_L or R_L) straight out of the shuttle display. H_X was a half pendulum in which writhing was absent. This behavior was sometimes hard to distinguish from the bird ending a display, since it is the presence of writhing that is particularly visually and acoustically salient. Thus, presence of H_X was likely underestimated. Another was a half pendulum with double writhing (H_W), in which the bird did an H with two instances of writhing rather than one, with a brief pause in between the first and second bout of writhing. In rare occasions, a bird briefly paused sound production mid-writhing to avoid colliding with an obstacle (such as a branch). These were coded as normal writhing; Hw specifically referred to a bird which ceased writhing, then resumed it when no external circumstance (such as an obstacle) might have influenced the bird. Moreover, H_W differed from the rare sequence 'HH' in that in 'HH' the male returned to the female after the first H, before the second. Another way that half-pendulum writhing varied was in the number of chirrups in a single bout of writhing, which, rather than considering this a variant, we instead treated separately as a quantitative character (presented below) (Table 3).

Pendulums (P). One apparent pendulum display was recorded from one Rufous Hummingbird in Alaska; this display was not visually observed by the recordist. Most hybrids expressed normal pendulum displays; hybrids that were *entirely* missing the pendulum display were always on the Rufous side of the hybrid zone and phenotypically indistinguishable from Rufous Hummingbird.

There were three variants of P that were performed by a few hybrids (Figure 3A, Table 2), including: P_D, two pendulums in the same direction, i.e. the element ‘turns around’ was missing at the end of the first pendulum; pendulums missing writhing (P_X), and pendulums with ‘double writhing’ (sequence *abc$\underline{bc}$g*), which meant that at the end of the display, the bird inserted a second bout of writhing (underline), with a short break in between the first and second bout, similar to half pendulums with double writhing (described above).

Allen's dives

Allen’s dives (A) included 'writhing on the ascent', i.e. as the male ascended for the dive, he writhed and produced chirrup sounds. Functionally, the inclusion of writhing on the ascent means that the start of a dive (in which the male descended slightly then ascended with writhing) is essentially a pendulum display that is incorporated into the beginning a dive.

There were six display variants of A: A_A, A_X, A_M, $A_{L,}$ A_{LX}, and A_{LA}. A_A was an aborted dive (N = 51 birds): the male ascended for a dive but then aborted it halfway through. In nearly all instances, the dive was aborted when the male simply flew away, or landed on a perch, ending the bout of display, omitting element sequence *ebcg* that is the second half of the dive (Figure 3, Table 2). Sometimes dives were aborted in response to an intruding male (e.g. the focal male would break off displaying to chase the other male; or the intruding male would attack the displaying male), but it was also common for the dive to be aborted without a clear external reason.

In A_X, the male omitted writhing at the end of the dive. This was the sole behavioral character that showed a statistically significant difference between the subspecies of Allen's Hummingbirds (Figure 4A); and it did not vary between Allen's and Rufous hummingbirds.

Dive variant A_L (N = 12 birds) was an Allen's dive missing the initial descent, such that the dive stated with a long ascent. This variant occurred most often when a bird ascended for a dive straight out of a shuttle display: a regular A dive, except lacking the entire 'incorporated pendulum' that begins a typical Allen's dive (Figure 3, Table 2). This variant was the result of an element sequence that mapped to two different display sequences, which are shown in Figure S1 A-C. Specifically, the sequence S*bcdgebcg,* can be interpreted as either display sequence SA_L or equivalently SHR_L (Figure S1A). Much more rarely, birds performed the sequence S*<u>bcbc</u>dgebcg*, i.e. the bird performed writhing twice in a row (repeated *<u>bcbc</u>*).

This dive type (including R_L, below) could be construed as atavistic: the outgroups Calliope Hummingbird (*S. calliope*) and Broad-tailed Hummingbird (*S. platycercus*) normally begin their dives by ascending straight up out of a shuttle display (Clark et al. 2018), and this is not normal for either Allen's or Rufous. A total of 23 hybrids (7% of hybrids) performed at least one A_L or R_L -type dives, as did six rufous (13%) and one Allen's (Table 2).

A_D (N = 1 bird) was an Allen's dive that lacked 'turn around' at the end of the dive. This meant that it ascended for another dive (of type A_L) without turning around at the end of the previous dive (Figure S1E). Although extremely rare in this dataset, this is precisely how the two closest outgroups within *Selasphorus* dive: both Calliope (*S. calliope*) and Broad-tailed Hummingbird (*S. platycercus*) ascend for a subsequent dive out of the previous dive, without turning around prior to the ascent. That is, out of 5,283 dives recorded in the present study, this one individual performed an U-shaped dive under the terminology of Clark et al. (2018). A_{LA} dives combined A_L ascent and were also aborted. A_{LX} was of A_L type ascent and also was missing writhing at the end of the dive (Figure 3).

A_M was the Allen's 'mini-dive', performed by N = 3 birds. This display was intermediate between a regular pendulum display and a regular dive. In essence, it appeared that a bird began to ascend for a dive, then halfway up, converted what was going to be a dive into a large pendulum. We treated this display a variant of the dive, rather than a pendulum, because the start of this display was split into A-type and R-type, like dives (see below). After ascending, the bird turned and descended but produced a faint or inaudible dive sound (element sequence: *abcd'gebcg*). We believe birds that performed this display still spread their tail (which is a component of element *e*) in at least in some instances, because the tail-generated portion of the dive sound was faintly audible in some recordings of mini-dives. Likely it was faintly audible because the bird's speed was reduced relative to a normal dive. We first noticed this display type in birds sampled in 2019, and suspect that in prior years we sampled additional birds that performed this display but mistakenly coded them as a more common display type (likely, P).

Rufous dives

The same dive variants that occurred in Allen's-type dives also occurred in Rufous-type dives, which included R_A, R_X, R_M, R_L and R_{LA} (Figure 3, Table 2). These had the same element sequences as described above for A-type dives, except the initial ascent for the dive lacked writhing (this was the kinematic difference between A and R dives). Finally, when a bird aborted a dive, the bout nearly always also ended, but two male Rufous from one population near the hybrid zone sometimes performed R_A repeatedly in sequence; they ascended for a dive, but then descended straight back to the female without following the normal kinematics or sounds of a dive, arrived at the female, then turned and ascended again for another aborted dive (R_{LA}). One bird, E08526, from the hybrid zone, was an outlier in that most of his dives were $R_{M.}$

Display variant prevalence

Display variant prevalence in Rufous, Allen's and hybrids is shown in Figure 4. Sedentary Allen's differed from migratory Allen's in one respect: sedentary Allen's were much more likely to omit writhing at the end of the dive ($p < 0.001$) than were migratory Allen's, hybrids, or rufous (Figure 4A). Apart from this we detected no differences between the two subspecies of Allen's Hummingbird, and remaining analyses presented here lump these two subspecies together. Some display variants seemed more prevalent within the hybrid zone: pendulums and half pendulums without writhing (Figure 4C), 'mini-dives' (Fig. 4D), and double writhing in half pendulums or pendulums (Fig. 4E). Note that as there are approximately 10 times as many hybrids sampled (N = 306) as sampled for either parental species (N = 35, 46), thus more outliers are expected among the hybrids due to chance alone (Figure 4).

The principal components analysis of display type prevalence in the collective dataset of birds, normalized for per bout (i.e., accounting for different males having different numbers of bouts sampled), are shown in Figure 4F with factor loadings in 4G. The first five PCs together accounted for 40.7% of the variation. PC 1 (12%) loaded on Allen's vs Rufous display types (Fig. 4C), while additional PCs loaded on variant display types as well as the shuttle display. PC2 loaded most strongly on variants H_D, A_M and A_L, for instance. Principal components above PC2 loaded on other combinations of variants (not shown). Many hybrids sat in between Allen's and Rufous Hummingbirds (green and orange polygons in Fig. 4F).

Chirrups

Chirrups (Figure 5) occurred in four display types: in half pendulums (all populations), at the end of the dive (except A_X and R_X dives; all populations); in pendulum displays (Allen's and hybrids); and in the ascent for A-type dives (Allen's and hybrids). Chirrup number was highly variable in certain hybrids during dive and half pendulum displays. Chirrups in ascents for Allen's dives never differed from chirrups in pendulum displays (data not shown), suggesting that the initial part of A dives is homologous to a regular pendulum display. By contrast, mean chirrup counts in the half pendulum, pendulum, and at the end of the dive varied from each other and varied across the hybrid zone (Figure 5 D-F), implying that these are distinct displays. Some hybrids were transgressive for chirrup number in the half pendulum display and the pendulum display (Fig. 5E, F).

Display bouts

We recorded a grand total of 4,399 display bouts. The median male had 11 bouts recorded; 12 birds had < 6 bouts, while 277 birds had between 10 and 15 bouts recorded. One outlier hybrid, SDNHM56915 performed 71 bouts over the course of 3 hours (before acclimating to the caged female), and two additional male Allen's (SDNHM 56918 and 56920) had 66 and 75 natural bouts recorded (performed to wild stimuli), each recorded over the course of > 1 month.

Within bouts, the number of displays varied from 1 to 29, with an approximately lognormal distribution (Figure 6). Allen's hummingbirds performed 6.3 ± 1.9 (mean ± sd) displays per bout, hybrids performed 4.8 ± 2.2 displays per bout, and Rufous hummingbirds, 2.6 ± 0.7 displays per bout. Bout length was driven by sequences of pendulums (many in Allen's, intermediate in hybrids, none in rufous). In both species and (apparently) all hybrids, shuttle displays were often performed as a single bout of display, either with or without a subsequent half pendulum (H) appended to them. In Allen's, display bouts nearly always began with either a pendulum display or a shuttle display, often followed by pendulums. Although a couple individuals did rarely perform a bout that was just a dive (without preceding pendulum displays), and a few Rufous type dives were recorded, both of these displays were rare in Allen's (Table 2). In Rufous, display bouts began either with a shuttle display or with dives, and only one putative pendulum display was recorded from a Rufous.

k-mer analysis of display sequences

There was substantial variation in display sequences. For the k-mer analysis we settled on k = 3, because a sequence length of 3 captured a substantial amount of the sequence variation that clearly differed between the species. The full k-mer alphabet including all variants contained 24 letters (Table 2). This full alphabet had a total of 24^3 = 13,824 possible 3-mers, but we do not present results on this alphabet, because the differences between the 6-letter alphabet and the 24-letter alphabet did not materially affect our results or conclusions, and it was easier to present analyses on an alphabet of six, with 216 rather than 13,824 possible 3-mers.

Out of a grand total of 29,693 occurrences of 129 types of 3-mers, PPP was the most prevalent (n = 6,384), and an 11 additional 3-mers occurred more than 700 times each (Figure 7A). Figure 7B shows a Venn diagram of the 73 3-mers that had a prevalence of >0.1% in at least one group (Allen's, Rufous, or hybrid), and table S1 gives a complete summary of the data by 3-mer and category.

Common 3-mers of Allen’s Hummingbird included EPP, PPP, PPA, PAE, and SHP (Figure 7, Table S1). Defining attributes of this phenotype included the presence and repetition of P, starting displays with P, the presence of A (Allen’s type dives), which almost always followed P, the near absence of 3-mers that included AA (two dives in a row), and the complete absence of the 3-mer AAA (three dives in a row). While most Allen's Hummingbirds only performed A type dives, as we did record a few R type dives (Table 2), REE was above 0.1% within Allen's and as a result, scored as a 'universal' 3-mer in Fig 7B, rather than as ‘hybrid-rufous’. In this respect the depiction in Figure 7 is slightly misleading: this 3-mer was common in Rufous and hybrids and rare in Allen’s (Table S1).

Common 3-mers of Rufous Hummingbird included ERR, RRR, RRE, and SHR (Figure 7, Table S1). Defining attributes of this phenotype included the repetition of R, the initiation of bouts of display with an R (e.g. ERR), and the lack of pendulum displays.

All hybrids incorporated parental displays into their repertoires and many birds categorized as “hybrids” were phenotypically indistinguishable from Allen’s or Rufous. However, many other hybrids included sequences that were nearly (prevalence of < 0.1%) or entirely nonexistent (0.0%) in the parent species (“Hybrid-only 3-mers” in Fig. 7). In particular, many hybrids frequently performed sequences that included both R (Rufous dives) and P (Figure 7C), a combination that was nearly nonexistent in the parental forms. Other distinguishing hybrid 3-mers consisted of a single P within a 3-mer (e.g., PRR); Allen’s performed single pendulums, but fairly rarely. Another was a single R within a 3-mer (e.g., PPR); 3-mers that included both A and R type dives (e.g., PRA) (Figure 7, Table S1), terminating a bout with a single P (e.g., RPE). That is, rather than performing multiple pendulum displays and then ending with a single dive (as Allen’s hummingbird often does) instead the hybrid performed one or multiple dives and then ended with a single pendulum display, inverting the order of these displays relative to Allen’s Hummingbird.

Figure 8 shows the principal components (PC) analysis of 3-mers. PC 1 loaded strongly on an Allen's – Rufous display axis (10.4% of variation), with rufous-like 3-mers (RRR, REE, EER, ERR) on one side, and Allen's 3-mers (EEP, EPP, PPP, AEE, PAE, PPA) on the other. We had naively predicted multiple clusters within 3-mer PC space, but in fact all hybrid 3-mers (common in some hybrids but rare or absent in Allen's and Rufous) loaded positively together in PC2 (4.8% of variation). Principal components above the second one loaded on 3-mers in ways that were difficult to apply a biological interpretation, or loaded strongly on a single individual bird.

3-mer repertoire

The measured repertoire size of individual birds varied from 6 to 69 different 3-mers. In Allen's we recorded 10 to 29 3-mers (mean: 17.0), Rufous expressed 6 to 23 3-mers (mean: 12.6), and hybrids expressed 8 to 69 3-mers (mean: 20.4). Accumulation curves of 3-mer repertoire as a function of the recorded bouts of display indicate that the behavioral repertoire was under-sampled (Figure 9A). Our intended sample size of 10 bouts per bird did not adequately sample the complete repertoire of many of the birds. A greater number of 3-mers were recorded in birds with a greater number of sampled display bouts (Figure 9B), but in addition, after accounting for variation in the number of bouts sampled, hybrids tended to express more 3-mers than Allen's or Rufous (ANOVA, 'species' as factor, # bouts sampled as a covariate, $p < 0.0001$; excludes the 3 outliers that had >30 bouts sampled shown in 9A). Out of 387 birds, 93 birds fell above the 1:1 line in Figure 9C, meaning that they had more singleton 3-mers (3-mers sampled exactly once within all bouts recorded) than doubleton 3-mers (3-mers sampled two or more times). Since methods to estimate 'true' repertoire size from a finite sample work poorly when there are more singletons than doubletons (Botero et al. 2008), we did not attempt to estimate the 'true' repertoire size from the accumulation curves.

DISCUSSION

Here we have described the variation in courtship displays in Allen's Hummingbird, Rufous Hummingbird, and birds from a hybrid zone between the two. The behavioral phenotype is organized hierarchically (Barlow 1968). We considered four levels within the phenotype. *Kinematic elements* were operationally defined subunits that comprise the displays (Figure 1). The next level was *displays*, which were fixed action patterns (since they could be performed in isolation). These displays were performed in sequences, forming *bouts*. Finally, the highest level was the *repertoire*, the cumulative set of behaviors observed spanning all the bouts we recorded for an individual, which we specifically presented as the set of 3-mers performed by individuals.

Of these four levels, Allen's, Rufous, and their hybrids all employed the same kinematic elements (with only one small exception in the rare 'mini-dive' display). One reason for this near-uniformity at the lowest level is almost certainly that the kinematic elements were operationally defined, and thus our definitions were, by their nature, somewhat loose. Had we been able to record displays with more precision (such as high speed video to visualize individual wingbeats; or with a method that allowed us to quantitatively record the height of dives, rather than estimating by eye), it is entirely possible that we would have uncovered additional noteworthy variation at this level. But the fact that our definitions of kinematic elements were operational cannot be the only reason we found little variation at this level. *A priori* we predicted that that we would find occasional rare individuals that were missing certain elements that both parents had, such as writhing, and we also predicted we would find atavistic behaviors, behaviors absent in both parent species but present in an outgroup. These hypotheses were largely unsupported. Among 306 hybrids, we did not detect any individuals that were entirely missing any elements. Regarding atavistic behaviors, out of our sample of a grand total of 20,901 displays, we did observe a single bird perform a U-shaped dive, the same type of dive performed by the outgroups (*Selasphorus calliope* and *S. platycercus*), similar to a pattern Lorenz (1958) reported for displays in duck hybrids. Moreover, 23 (7%) of hybrids sometimes ascended for dives in the same way that the outgroups do (A_L and R_L type dives), but as 13% of allopatric Rufous and 3% of allopatric Allen's expressed this dive type at least once as well, whether this type of dive is actually atavistic is unclear.

In contrast to this relative uniformity at the lowest level of behavioral organization, the courtship *displays*, the *bouts*, and the *repertoires* varied substantially. That is, the variation was not in the elements themselves, but rather, how they were sequenced: by assembling the sequences in different orders, Allen's and Rufous Hummingbird performed materially different behaviors, as did many hybrids.

Display variants and transgressive segregation

Although we consider the displays we described to be fixed action patterns, like all FAPs, they were not entirely fixed (Barlow 1977): We detected a number of display variants (Figure 3). In most cases the variants comprised insertions or omissions of one or more kinematic elements relative to the 'typical' form of the display. Some of the variants we describe were rare (such as pendulum displays missing writhing, Figure 4C).

As we hypothesized, hybrids that expressed some of these variants had transgressive phenotypes, i.e. phenotypes that fall outside the range of variation of the parental forms due to recombination and the interaction of genes that are responsible for different traits (Hegarty 2012). Most transgressive phenotypes were related to the expression of the element 'writhing': some individuals omitted writhing in displays that normally have it (Figure 4C), or performed writhing twice in succession, rather than once (Figure 4E). Another form of transgressive writhing was within a single bout of writhing, in which certain individuals performed writhing for longer (more chirrups) than either parental species (Figure 5E, F). Since one of the major differences between Allen's and Rufous Hummingbird display phenotypes is that Allen's performs writhing when ascending for a dive, and Rufous does not (Figure 1), we hypothesize that these transgressive phenotypes arise as the result of complementary gene combinations that control expression of this phenotypic difference between the two species (Lexer et al. 2003).

Most examples of transgressive characters come from plants, such as two short plants that, when crossed, give rise to a taller hybrid (Rieseberg, Archer & Wayne 1999). Transgressive segregation of behaviors has been previously described from crosses of two *Mus musculus* subspecies (Hiadlovská et al. 2012), as well as duck displays (Lorenz 1958; Sharpe & Johnsgard 1966) and two hybrid hummingbird courtship displays of close relatives of the species studied here (Clark, Feo & Bryan 2012; Wells, Bradley & Baptista 1978). While the duck studies were of controlled lab crosses that produced a sample size of F2s, these prior hummingbird examples describe wild F1 hybrids and had small sample size. The allegedly transgressive behavior reported by Wells et al. (1978) for Anna's × Costa's Hummingbird (*Calypte anna* × *C. costae*) hybrids (of flying in a large horizontal circle at the apex of a dive) has been subsequently observed occasionally in *C. costae* (CJC, pers obvs), and a behavior expressed during an ascent for a dive that Clark et al. (2012) suggested was transgressive in a Black-chinned × Broad-tailed (*Archilochus alexandri* × *S. platycercus*) hybrid has subsequently been observed in Broad-tailed (Simpson & McGraw 2018; Simpson pers. comm.), suggesting that neither of these prior examples are actually transgressive. In contrast to these two examples, the transgressive behaviors of the *Selasphorus* hybrids we describe here (Figures 4, 5) do not suffer from small sample size and inadequately described parental display behaviors.

Hybrid sequence variation

One general finding was that courtship displays of some hybrids (especially, those near the middle of the hybrid zone) were more variable than either parent species. The presence of display variants (described above) in hybrids was a component of this variation, but this heightened variation was even more prevalent in the 3-mer repertoires of hybrids (compare Figure 4F against Figure 8). The most variable hybrids were striking: in one bout of display they might perform a sequence that resembled an Allen's Hummingbird display, performing several pendulums followed by a dive; then in the next bout of display, perform a sequence of several dives in a row, like Rufous hummingbird (all such birds also performed many sequences that included hybrid-only 3-mers shown in Figure 7B, 7C). While it is already well-known that hybrids (e.g. F2 backcrosses) tend to exhibit high variation, the within-individual aspect of this variation we describe here is noteworthy, because the nature of this variation differs from how variation in other types of traits is expressed.

F2 backcrosses and other hybrids are famously variable in trait expression *as populations* because different F2 individuals contain variable combinations of independently segregating genes inherited from their parent taxa. Yet, F2 *individuals* are usually not variable. One reason for this is trivial: most animal morphological characters cannot vary within an individual because each animal only has one iteration of the trait: a hybrid animal has only one left foreleg or head, for instance. The type of trait that could have high intraindividual variation are modular traits that are highly repeated, like hairs, feathers, or flowers. Of these, hybrid flower morphology has been well-studied. Flowers can show high among-individual variation in F2 crosses or hybrid swarms but in F2 crosses, flowers apparently do not show substantial morphological variation within an individual (Tavares et al. 2018). For instance, Alexandre et al. (2015) generated F2 backcrosses of two *Rhytidophyllum* species, which yielded high inter-individual variation in floral morphology (see their figure 1). But this population variation is not recapitulated in the morphology of multiple flowers from the *same individual* plant: the flowers from a given plant have low variation. The only other example we could find of a trait like this was also behavioral: an individual hybrid *Sayornis* phoebe sang variable songs that mixed components from both parent species (McCallum & Pieplow 2010).

The high within-individual behavioral variability in hybrids, uncovered here, is intriguing. Presumably it is the product of how genes from the parent species interact to produce the specific neural circuit(s) in the brain that give rise to the courtship displays. Since the ways that the cerebellum produces flight (let alone flight displays) remains unresolved in birds (D. Altshuler, pers. comm.), there appears to be little we can predict about the precise nature of how gene interactions would give rise to neural circuits that generate substantial within-individual variation.

This within-individual variation also made it harder to precisely characterize the phenotype of a given individual hybrid. For instance, while 10 bouts of display were not sufficient to fully sample the repertoire of any given bird, this was especially true of hybrids (Figure 9). One hybrid, SDNHM56915, was unusually motivated to display to a caged female and did not acclimate until after performing 71 bouts of display (in one day). This bird continued to include new 3-mers in his repertoire as we sampled additional bouts (Figure 9). Because all of our birds are under-sampled, it is difficult to resolve true fine-scale behavioral differences between individuals.

Are these displays really under genetic control?

Above we have implicitly assumed that the *Selasphorus* displays are predominantly under genetic control. Araya-Salas et al. (2019) provided evidence of social learning of displays in hermit hummingbirds (*Phaethornis*). Within hummingbirds, *Phaethornis* and *Selasphorus* are distantly related, and *Phaethornis* species perform different displays than the displays described here. For instance, hermits do not perform a dive. The effect size of the socially learned component reported by Araya-Salas and Wright (2019) of hermit displays was small; most of the variation in hermit displays that they measured cannot be attributed to social learning. While we cannot rule out subtle effects of social learning on the behaviors described here, several patterns are inconsistent with social learning being the primary mode of acquisition of displays. Transgressive behaviors are not expected under social learning. Instead, hybrids are predicted to copy (and therefore resemble) the displays of one parent species or the other. Copy errors could generate transgressive display types but if so these are expected to cluster geographically (akin to vocal dialects), which the variants described here did not do. There are two disjunct and independent parts of the zone: there is an east-west arm of the hybrid zone along the Klamath River in Siskyou Co, California, and separately, a north-south hybrid zone centered along the Pacific coast of Oregon. The centers of our two transects (roughly, Coos Bay, Oregon and Happy Camp, California) are about 170 kilometers apart, and similar display variants were found in both. Another line of evidence that these displays are not socially learned: Allen's Hummingbirds raised in isolation subsequently produced pendulum displays (Isenberg 1962). Moreover, F1 hybrids of various crosses of close relatives of *Selasphorus* tend to produce displays that are mosaics of the displays of parent species, rather than matching one parent or the other (Clark et al. 2012; Clark, Feo & van Dongen 2013; Wells et al. 1978).

Display function

We infer these displays are courtship displays, as males ardently performed them for females, and females do not perform any of the displays describe here, suggesting they have primarily evolved under the influence of female choice. However, we lack data on female choice so we do not know which aspects of the displays described here females may like or dislike.

The shuttle display seemed to serve both courtship and agonistic functions. Shuttle displays were frequently given to intruders, following the pattern that in many hummingbirds, aggressive displays and courtship displays can be difficult to differentiate (Mobbs 1982; Skutch 1973). Shuttle display length varied greatly in length, and while males sometimes directed long displays to the caged female because she could not evade the male, they also often directed short shuttle displays to the caged female, especially as a male became acclimated to the caged female and soon ceased displaying entirely.

k-mer sequence analysis

Our use of k-mer statistics to describe sequence variation appears to be the first use of this type of statistic for a behavioral sequence problem. A 3-mer is equivalent to a 3rd order Markov model, except instead of calculating a transition matrix of the probability of transitioning from one 3rd order model to another (i.e., the point of a Markov model), we simply counted the presence/absence of each occurrence of the 3-mer within the repertoire of each animal, and conducted descriptive statistics on these counts. Given that a large number of the 3-mers we detected were rare within the sample (Figure 7), and that our sequence repertoires were under-sampled (Figure 9), any attempt to calculate transition probabilities seemed likely to be error-prone.

Our choice of k = 3 means that our method was blind to variation in sequences that was higher order (longer length) than 3. We suspect that, at least in Allen's Hummingbird, there may have been some meaningful sequence variation at longer lengths. We say this because some males seemed to do a somewhat consistent number (up to ~15) of pendulum displays before switching to a dive, such that the number of repeats of P did not seem entirely random (although testing this formally would likely require recording more sequences from individual Allen's). But as hybrids appeared to be more random in their sequences, k = 3 seemed sufficient to capture much of the sequence variation, while keeping the sequences short enough to be manageable.

We initially attempted to use k-mers to describe *element* sequences, in order to use cluster analysis (of k-mers) to neutrally infer which sequences of elements were best considered natural units that we could consider to be displays, as opposed to the method we ultimately did use, of defining displays 'by eye'. The reason it was not feasible to use k-mers to identify displays from element sequences was that dives were typically 6 to 9 elements long (Table 2), meaning we would need to consider element sequences substantially longer than k = 9 in order to assess this statistically. Given that the number of possible sequences rises exponentially with k, we did not determine a way to identify element sequence clusters (for, say, k = 20) with an algorithm, and instead resorted to identifying natural element sequences by eye. As genomes can be assembled with high values of k (e.g. k = 90) (Sameith, Roscito & Hiller 2017), perhaps similar techniques can be adapted for use on behavioral sequences. The method we used here, in which we constructed a k-count matrix, is not amenable to k = 90. Applied to N = 4 letters, k = 90 would require a k-count matrix with $4^{90} = 1.5 \times 10^{54}$ columns, which is too many columns for a statistics program such as R.

For the hummingbird displays described here, we think that identifying displays from element sequences by eye has nevertheless yielded sensible results, for a specific reason: by and large, each element we have defined is done by the whole animal, such that the whole animal can only do one kinematic element at a time. For displays where different kinematic elements of the display can be produced independently, such as jumping spider displays where different legs move independently (Elias et al. 2012), defining displays 'by eye' may be problematic, and more sophisticated (algorithmic) methods would need to be developed instead.

Conclusions

Here we have used k-mer analysis to describe variation in male courtship display sequences of 35 allopatric Allen's Hummingbirds, 46 allopatric Rufous hummingbirds, and 306 birds from the hybrid zone between these two species. The lowest level of display, the 'kinematic elements', showed almost no variation among groups; all individuals perofmred displays comprising the same set of elements, with only one minor exception, the 'mini-dive' display performed by 6 hybrids. Instead the courtship differences between the species were at the level of displays. Many hybrids performed occasional display variants not observed in the parental forms, including transgressive displays and possibly atavistic displays. The parental forms and hybrids also varied in how displays were sequenced in bouts. Our use of 3-mers to characterize display sequences seemed to successfully capture much of the sequence variation observed. The full set of 3-mers each bird performed comprised the sequence repertoire, which also varied among the three groups. Some hybrids (particularly, from the middle of the hybrid zone) were characterized by having high variation in the sequences of displays, and thus, had large 3-mer repertoire size. This high within-individual variation of hybrid phenotype may be a feature unique to behavior.

REFERENCES

Alexandre, H., Vrignaud, J., Mangin, B. & Joly, S. 2015. Genetic architecture of pollination syndrome transition between hummingbird-specialist and generalist species in the genus *Rhytidophyllum* (Gesneriaceae). *PeerJ*, **e1028**.

Araya-Salas, M., Smith-Vidaurre, G., Mennill Daniel, J., González-Gómez Paulina, L., Cahill, J. & Wright Timothy, F. 2019. Social group signatures in hummingbird displays provide evidence of co-occurrence of vocal and visual learning. *Proceedings of the Royal Society B: Biological Sciences*, **286**, 20190666.

Baerends, G. P. 1958. Comparative methods and the concept of homology in the study of behaviour. *Archives Néerlandische Zoologie Supplement 1*, **13**, 401-417.

Barlow, G. 1977. Modal action patterns. In: *How Animals Communicate* (Ed. by T. Sebeok), pp. 98-134: Indiana University Press.

Barlow, G. W. 1968. Ethological units of behavior. In: *The Central Nervous System and Fish Behavior* (Ed. by D. Ingle), pp. 217-232: University of Chicago Press.

Botero, C. A., Mudge, A. E., Koltz, A. M., Hochachka, W. M. & Vehrencamp, S. L. 2008. How reliable are the methods for estimating repertoire size? *Ethology : formerly Zeitschrift fur Tierpsychologie*, **114**, 1227-1238.

Charif, R. A., Waack, A. M. & Strickman, L. M. 2008. *Raven Pro 1.3 User's Manual*. Ithaca, NY: Cornell Laboratory of Ornithology.

Clark, C. J. 2009. Courtship dives of Anna's Hummingbird offer insights into flight performance limits. *Proceedings of the Royal Society of London B*, **276**, 3047-3052.

Clark, C. J. 2014. Harmonic hopping, and both punctuated and gradual evolution of acoustic characters in Selasphorus hummingbird tail feathers. *Plos One*, **9**, e93829.

Clark, C. J. 2016. Locomotion-Induced Sounds and Sonations: Mechanisms, Communication Function, and Relationship with Behavior. In: *Vertebrate Sound production and Acoustic Communication* (Ed. by R. A. Suthers & T. Fitch), pp. 83-117. New York: Springer Handbook of Auditory Research.

Clark, C. J., Feo, T. J. & Bryan, K. B. 2012. Courtship displays and sonations of a male Broad-tailed × Black-chinned Hummingbird hybrid. *Condor*, **114**, 329-340.

Clark, C. J., Feo, T. J. & van Dongen, W. 2013. Sounds and courtship displays of the Peruvian Sheartail, Chilean Woodstar, Oasis Hummingbird, and a hybrid male Peruvian Sheartail × Chilean Woodstar. *Condor*, **115**, 560-577.

Clark, C. J., McGuire, J. A., Bonaccorso, E., Berv, J. S. & Prum, R. O. 2018. Complex coevolution of wing, tail, and vocal sounds of courting male bee hummingbirds *Evolution*, **72**, 630-646.

Clark, C. J. & Mistick, E. A. 2018. Kinematic control of male Allen's Hummingbird wing trill over a range of flight speeds. *The Journal of Experimental Biology*, **221**, jeb173625.

Clark, C. J. & Mitchell, D. E. 2013. Allen's Hummingbird (*Selasphorus sasin*). In: *Birds of North America Online* (Ed. by A. Poole). Ithaca, NY: Cornell Lab of Ornithology.

Elias, D. O., Maddison, W. P., Peckmezian, C., Girard, M. B. & Mason, A. C. 2012. Orchestrating the score: complex multimodal courtship in the *Habronattus coecatus* group of *Habronattus* jumping spiders (Araneae: Salticidae). *Biological Journal of the Linnean Society*, **105**, 522-547.

Hegarty, M. J. 2012. Invasion of the hybrids. *Molecular ecology*, **21**, 4669-4671.

Henderson, E. C. & Brelsford, A. 2020. Genomic differentiation across the speciation continuum in three hummingbird species pairs. *BMC Evolutionary Biology*, **20**, 113.

Hiadlovská, Z., Bímová, B. V., Mikula, O., Piálek, J. & Macholán, M. 2012. Transgressive segregation in a behavioural trait? Explorative strategies in two house mouse subspecies and their hybrids. *Biological Journal of the Linnean Society*, **108**, 225-235.

Isenberg, A. H. 1962. The breeding of Allen's hummingbird. *Avicultural Magazine*, **68**, 86-88.

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., Huetz, C., Hughes, M., Hyland Bruno, J., Ilany, A., Jin, D. Z., Johnson, M., Ju, C., Karnowski, J., Lohr, B., Manser, M. B., McCowan, B., Mercado, E., 3rd, Narins, P. M., Piel, A., Rice, M., Salmi, R., Sasahara, K., Sayigh, L., Shiu, Y., Taylor, C., Vallejo, E. E., Waller, S. & Zamora-Gutierrez, V. 2016. Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biological reviews of the Cambridge Philosophical Society*, **91**, 13-52.

Lexer, C., Welch, M. E., Raymond, O. & Rieseberg, L. H. 2003. The origin of ecological divergence in *Helianthus paradoxus* (Asteraceae): Selection on transgressive characters in a novel hybrid habitat. *Evolution*, **57**, 1989-2000.

Lorenz, K. Z. 1941. *Comparative studies of the motor patterns of Anatinae*. Cambridge, MA: Harvard University Press.

Lorenz, K. Z. 1950. The comparative method in studying innate behavior patterns. *Symposia of the Society for Experimental Biology*, **4**, 221-254.

Lorenz, K. Z. 1958. The evolution of behavior. *Scientific American*, **199**, 67-77.

McCallum, D. A. & Pieplow, N. D. 2010. A reassessment of homologies in the vocal repertoires of phoebes. *Western Birds*, **41**, 26-43.

Mobbs, A. J. 1982. *Hummingbirds*. Hindhead, Surrey: Triplegate Ltd.

Myers, B., Rankin, D., Burns, K. J. & Clark, C. J. 2019. Behavioral and morphological evidence of an Allen's × Rufous Hummingbird (*Selasphorus sasin × S. rufus*) hybrid zone in southern Oregon and northern California. *The Auk: Ornithological Advances*, **136**, ukz049.

Myers, B. M., Burns, K. J., Clark, C. J. & Brelsford, A. 2021. The population genetics of nonmigratory Allen's Hummingbird (*Selasphorus sasin sedentarius*) following a recent mainland colonization. *Ecology and Evolution*, **11**, 1850-1865.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E. & Wagner, H. 2020. vegan: Community Ecology Package. *R package version 2.5-7*.

Patricelli, G. L., Krakauer, A. H. & Taff, C. C. 2016. Variable Signals in a Complex World: Shifting Views of Within-Individual Variability in Sexual Display Traits. *Advances in the Study of Behavior*, **48**, 1-67.

Rieseberg, L. H., Archer, M. A. & Wayne, R. K. 1999. Transgressive segregation, adaptation and speciation. *Heredity*, **83**, 363-372.

Sameith, K., Roscito, J. G. & Hiller, M. 2017. Iterative error correction of long sequencing reads maximizes accuracy and improves contig assembly. *Briefings in Bioinformatics*, **18**, 1-8.

Schleidt, W. M. 1974. How "fixed" is the fixed action pattern? *Zeitschrift für Tierpsychologie*, **36**, 184-211.

Scholes, E. 2006. Courtship ethology of Carola's parotia (*Parotia carolae*). *Auk*, **123**, 967-990.

Scholes, E. 2008. Evolution of the courtship phenotype in the bird of paradise genus *Parotia* (Aves: Paradisaeidae): homology, phylogeny, and modularity. *Biological Journal of the Linnean Society*, **94**, 491-504.

Sharpe, R. S. & Johnsgard, P. 1966. Inheritance of behavioral characters in F2 Mallard × Pintail (*Anas platyrhynchos L. × Anas acuta L.*) hybrids. *Behaviour*, **27**, 259-272.

Simpson, R. K. & McGraw, K. J. 2018. Two ways to display: male hummingbirds show different color-display tactics based on sun orientation. *Behavioral Ecology*, **29**, 637-648.

Skutch, A. F. 1973. *The Life of the Hummingbird*. New York: Vineyard Books, Inc.

Tavares, H., Whibley, A., Field, D. L., Bradley, D., Couchman, M., Copsey, L., Elleouet, J., Burrus, M., Andalo, C., Li, M., Li, Q., Xue, Y., Rebocho, A. B., Barton, N. H. & Coen, E. 2018. Selection and gene flow shape genomic islands that control floral guides. *Proceedings of the National Academy of Sciences*, **115**, 11006.

Team. 2019. RStudio: Integrated Development for R. Boston, MA, http://www.rstudio.com: RStudio, PBC.

Team, R. C. 2018. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Thorpe, W. H. 1951. The definition of some terms used in animal behaviour studies. *Bulletin of Animal Behaviour*, **1**, 34-40.

Wagner, G. P. 2014. *Homology, genes, and evolutionary innovation*. Princeton, New Jersey: Princeton University Press.

Wells, S., Bradley, R. & Baptista, L. F. 1978. Hybridization in *Calypte* hummingbirds. *The Auk*, **95**, 537-549.

Wenzel, J. W. 1992. Behavioral homology and phylogeny. *Annual Review of Ecology and Systematics*, **23**, 361-381.

Wilkinson, S. 2018. kmer: an R package for fast alignment-free clustering of biological sequences. *R package version 1.0.0*, **https://cran.r-project.org/package=kmer.**

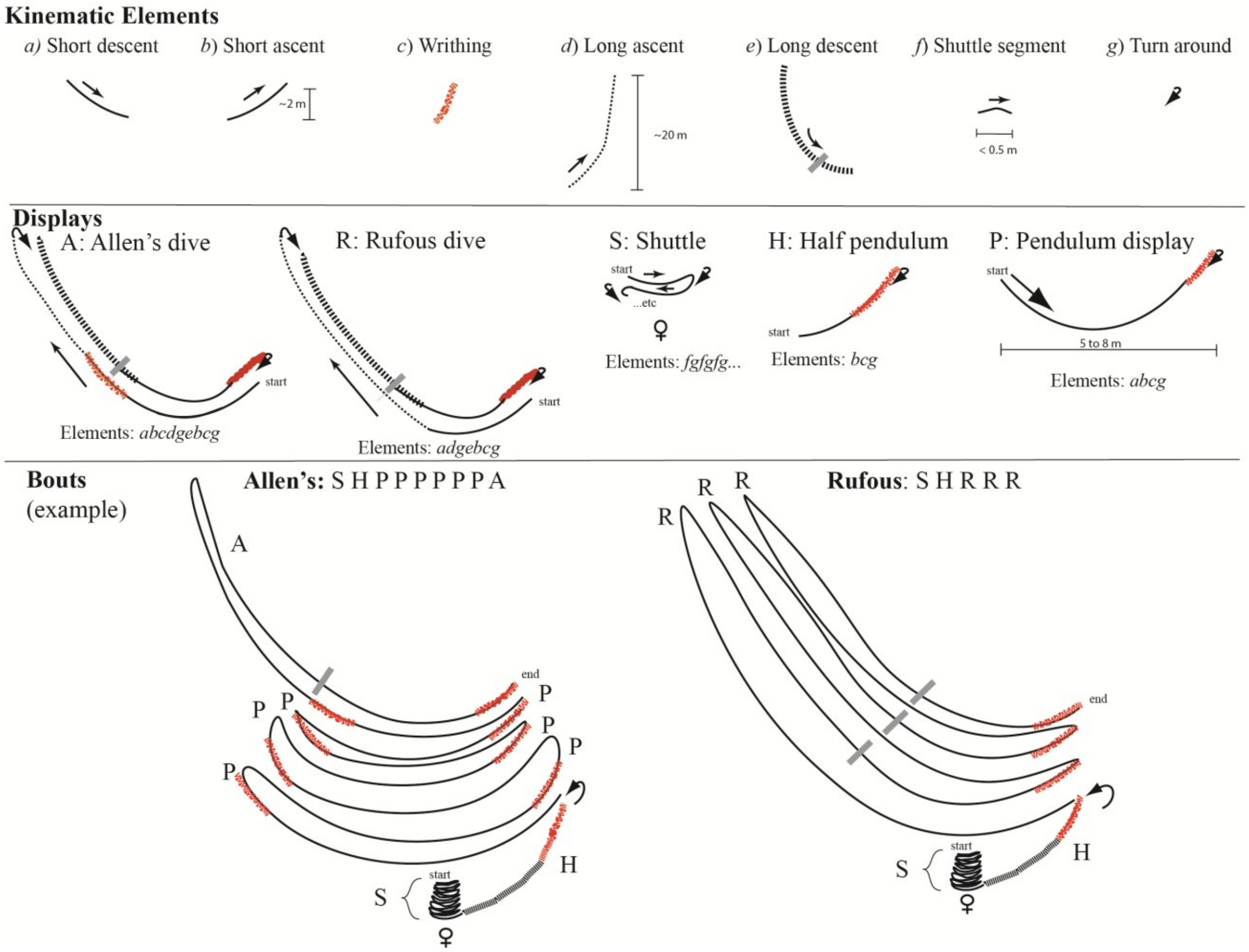

Figure 1. **Kinematic elements, displays, and example courtship bouts of Allen's and Rufous Hummingbird. Top:** We defined 7 kinematic elements that comprise the displays, symbolized with lowercase letters *a*-*g*; element f (shuttle segment) is omitted to simplify the analyses. Grey hashmark in *e* indicates dive sound made by tail. Writhing, *c*, occurs during element *b*. This notation follows Myers et al. (2019), except for element *g* added here. **Middle**: Courtship displays A, H, P, R and S. The only difference between the dives is Allen's dive (A) contains writhing on the ascent for the dive (red squiggly line), while the Rufous dive (R) does not. The shuttle (**S**) comprised any number of repeats of a short back-and-forth flight. **Bottom**: Example bouts of display typical to Allen's and Rufous, each beginning with a shuttle (S) and ending with a dive. Subsequent pendulums/dives have been drawn slightly offset; in reality, the male flies back and forth through the same airspace above the female repeatedly. Displays are not drawn to scale.

Figure 2. **k-mer analysis.** 10 display sequences (bouts) from a hypothetical Allen's hummingbird. The sequences were concatenated after placing dummy letters EE at the start and end of each bout. Applying k = 3 to this concatenated sequence of 111 letters yields 109 3-mers; the presence of EE (given k = 3) in between sequences ensures that the order in which bouts were concatenated does not affect the k-count matrix. This sequence contains 18 unique k-mers (first instance underlined). The most common sequence, 3-mer PPP occurred 58 times (green), while four 3-mers occur once (red), and others are underlined blue.

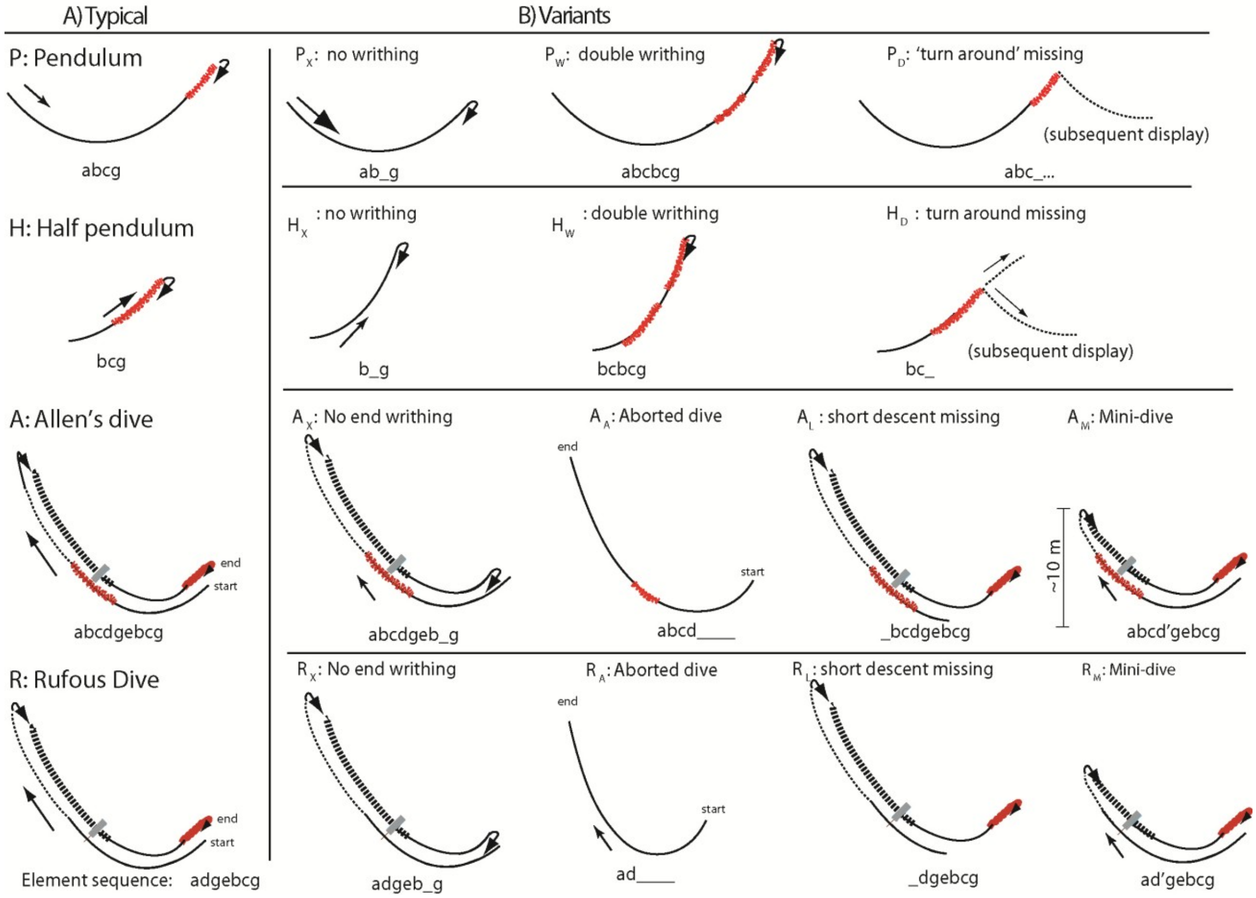

Figure 3. **Courtship displays recorded in this study**. See Table 2 for summary statistics. **A**) Typical versions of display, including element sequence (lower case letters, as defined in Fig. 1 and Table 1). **B**) Observed variants. Gaps in the typical sequence caused by missing elements indicated with '_'. Variants A_{LX}, $A_{LA,}$ R_{LA} and A_D, not shown. The first three are dives with two atypical features, e.g. A_{LX} dives both omit the initial ascent (subscript L) and also the end writhing (subscript X). A_D is shown in Fig. S1E, was a single dive in which the male ascended for another dive out of the end writhing of the previous dive. For the 'mini-dive' (A_M and R_M), element d' is about half the height of the normal element d.

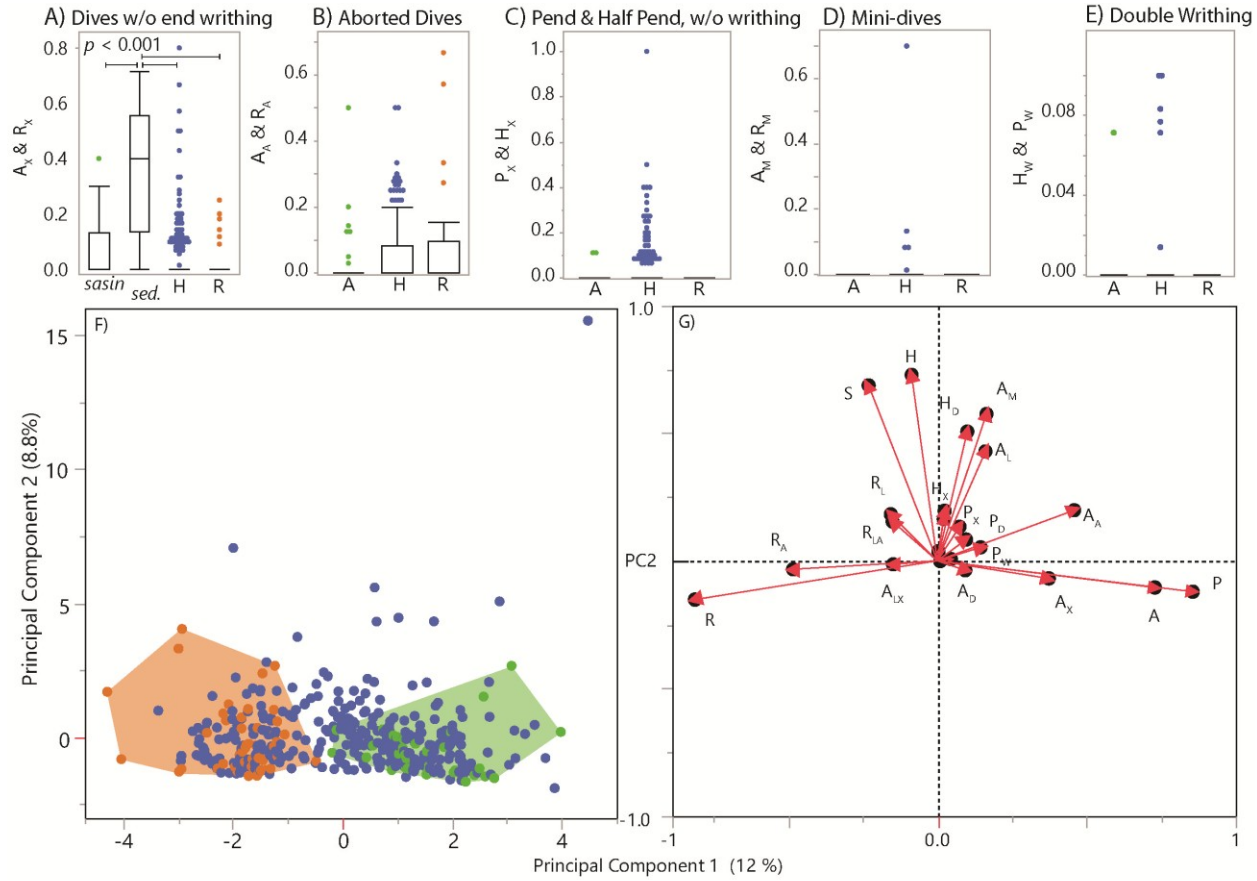

Figure 4. **Frequency of display variants across Allen's (A), Hybrids (H) and Rufous (R).** Statistics for A-E: ANOVA, followed by post-hoc Tukey tests. Data have been normalized per bout (# of times bird performed that display type divided by total number of bouts it performed, not including bouts that were only shuttle display) to control for variation in numbers of bouts recorded. Box and whiskers are 1 and 2 SD. Note that box and whiskers are 0 for all three categories in C, D, and E; all visible points are outliers. Sample size: 35 Allen's, 306 Hybrids, 46 Rufous. *S. s. sasin* (N = 26) did not differ from *S. s. sedentarius* (N = 9) and so data for the two subspecies are pooled in all panels except A.

A) Dives missing end writhing were more prevalent in sedentary Allen's hummingbird (*Selasphorus sasin sedentarius*, N = 9 individuals) than in *S. s. sasin* (N = 26), hybrids (N = 306), or Rufous (N = 46). **B**) Prevalence of aborted dives did not differ among groups (Tukey test, all $p > 0.2$) and this character was not transgressive. **C**) Pendulums and half pendulums missing writhing (median = 0) did not differ among populations (Tukey test, all $p > 0.05$) but there were 26 transgressive hybrids. **D**) Mean expression of the 'mini-dive' phenotype did not differ among populations (Tukey test, all $p > 0.8$); five hybrids performed mini-dives. **E**) Double writhing did not vary among populations (Tukey test, all $p > 0.6$); six hybrids and one Allen's performed displays with 'double writhing'.

F) Principal component analysis of display types for N = 386 male hummingbirds, normalized by the number of bouts each male performed. Polygons enclose phenotype space for allopatric Allen’s (green) and Rufous (orange). **G**) Factor loadings for the display types. PC 1 loaded heavily on Rufous (R, R_A, R_X) versus Allen's (A, P, A_X, A_A) display types. PC2 loaded on display variants as well as shuttle displays (S and H).

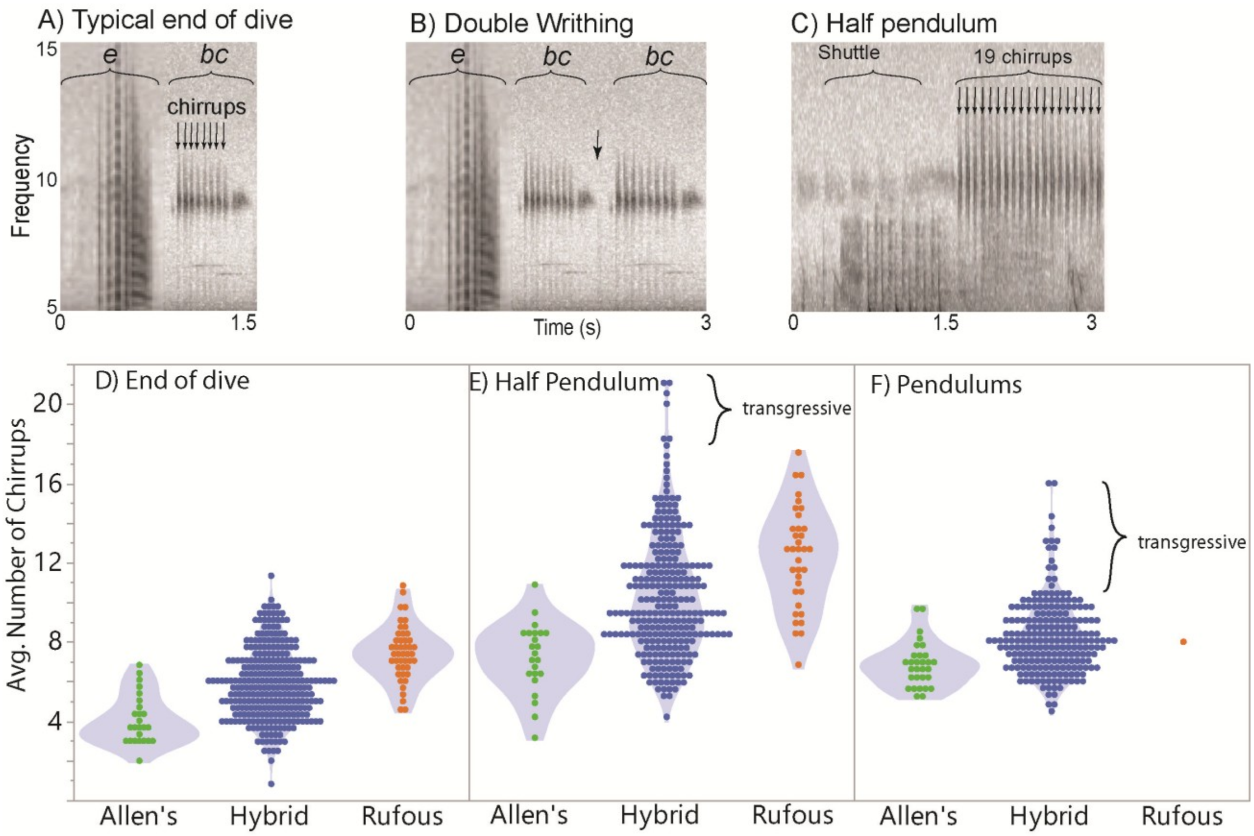

Figure 5. **Spectrograms and counts of chirrup sounds produced by writhing during displays**. Writhing (red squiggly line in Figs 1 and 3) occurred in pendulum displays, half pendulums, at the beginning of A dives, and at the end of both A and R dives. **A**) Typical writhing at the end of a dive (elements *b* and *c*). Writhing at the end of the dive (elements *b* and *c*) occurred right after element *e* which included sounds made with the tail-feathers. During each writhing motion, one chirrup sound was made with the wings (8 arrows). **B)** In 'double writhing' (displays H_W and P_W) elements *bc* are repeated with a short break in between (arrow), during which time the male continues to fly up without writhing. **C)** Writhing/chirrup sounds during a half pendulum (19 chirrups shown). **D-F)** Number of chirrups at the end of dives, in half pendulums, and in pendulum displays. **D)** Number of chirrups at the end of the dive was significantly greater in Rufous than in Allen's (ANOVA, post-hoc Tukey test, $p < 0.0001$). **E)** Number of chirrups was significantly greater in Rufous than in Allen's (ANOVA, post-hoc Tukey test, $p < 0.0001$), and 7 hybrids were apparently transgressive. **F)** Rufous hummingbirds do not normally perform pendulum displays; 33 hybrids fell above range of chirrup values of Allen's Hummingbird and were apparently transgressive for chirrups. Values shown are averages from up to 6 displays per bird. Chirrups during the ascent for A dives were never different from chirrups in pendulum displays and so are not shown.

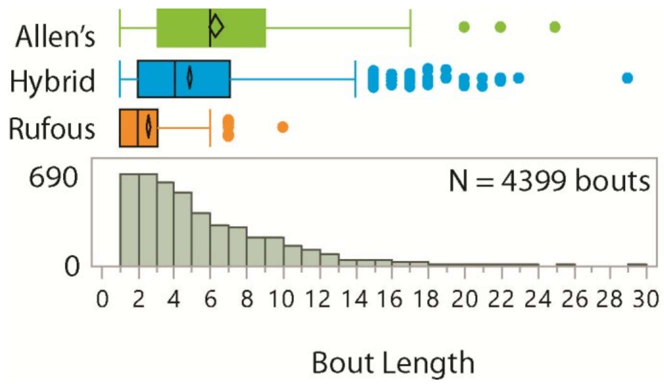

Figure 6. **Bout length varied from 1 to 29 displays across all 387 birds.** Bout length does not include dummy "EE" appended to the start and end of each bout (i.e. a single shuttle display has a length = 1 here, but is "EESEE" in the 3-mer analysis). Diamond: median, box: 25th-75th percentiles, whiskers: 5th-95th percentiles.

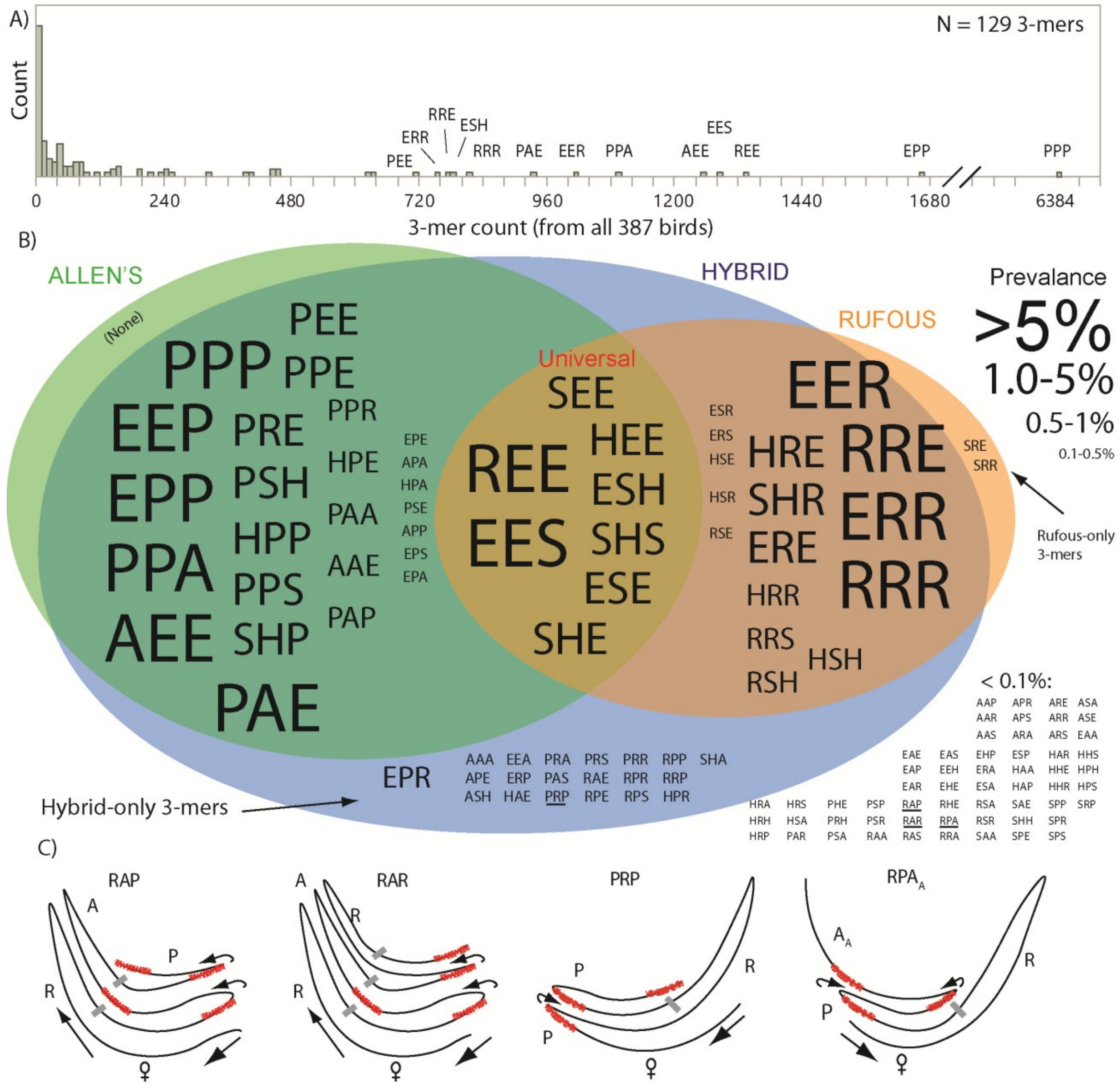

Figure 7. **3-mer counts among Allen's, Rufous, and hybrids. A)** Out of a grand total of 29,693 occurrences of 129 3-mers, a handful of them had high counts, especially PPP. **B) 3-mer prevalence in Allen's (green), Rufous (brown), and Hybrids (blue) of the 73 most common 3-mers.** An additional 58 3-mers with prevalence of < 0.1% in all 3 categories are shown in the lower right. 3-mers are sized according to the maximum % of total 3-mer count for each group (Allen's: N = 4,127, Hybrid N = 23,352, Rufous N = 2,114). For instance, RRE is 2.2% of 3-mers for hybrids and 11.2% for rufous, so is sized as having a prevalence of >5%. Source data are in Table S1. **C) Four examples of 3-mers that were only expressed in hybrids** (underlined in B).

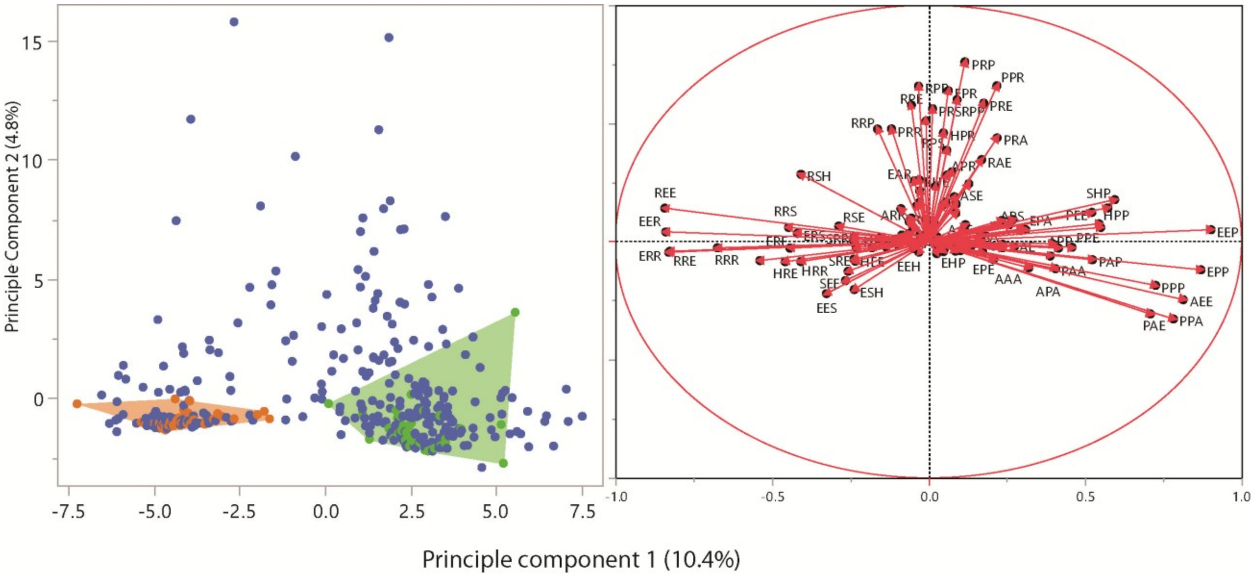

Figure 8. Principal components of 3-mers. PC1 loaded strongly on a Rufous-Allen's continuum, with Rufous-typical 3-mers such as RRR and RRE on one end, and Allen's-typical 3-mers such as PPP, PPA, and AEE on the other. PC 2 loaded strongly on 3-mers not present at all (e.g. PRP, EPR) in the parent species, (or rare, such as PPR). Orange dots and orange polygon are allopatric Rufous Hummingbirds; green dots and polygon are allopatric Allen's Hummingbird.

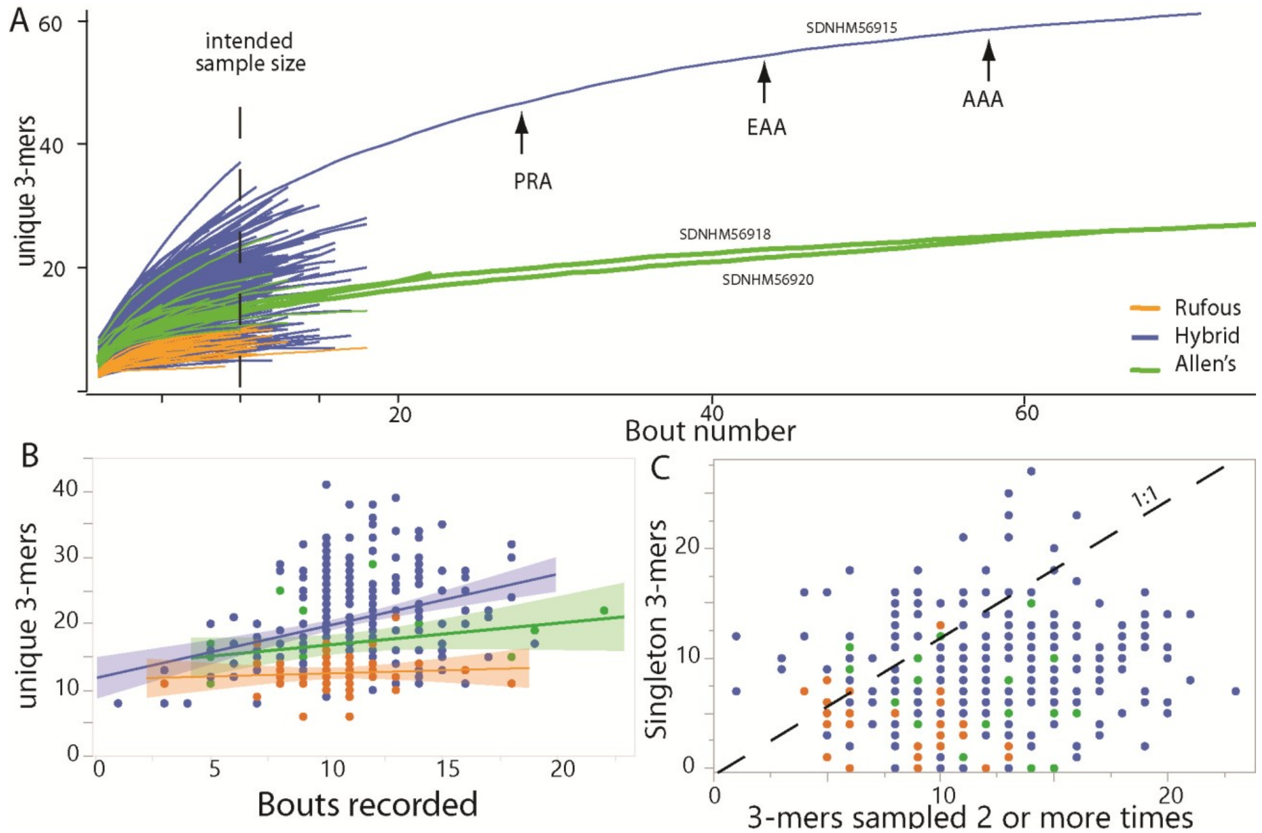

Figure 9. **The sequence repertoire (i.e., all 3-mers performed by each individual) was under-sampled. A**) Accumulation curves of 3-mer for Allen's (N = 35), Rufous (N = 46), and hybrids (N = 306). SDNHM56915 was an exceptional hybrid that performed bout after bout to a female in a cage over the course of a few hours, before acclimating after 71 bouts. 3-mer sequence 'PRA' first appeared in bout 28, 'EAA' in bout 44, and 'AAA' in bout 57. This bird had 14 singleton 3-mers and 55 3-mers sampled more than once at bout 71. SDNHM56918 and SDNHM56920 were recorded over the course of ~1 month (each), performing natural displays to wild females, rather than displays to a caged female. After 66 bouts, SDNHM56918 had 28 singletons and 14 3-mers performed more than once. **B**) Number of unique 3-mers sampled increased with the number of bouts of display sampled. **C**) Ninety four individuals had more singleton 3-mers (3-mers expressed exactly once in all of the bouts recorded), relative to the number of 3-mers expressed two or more times within all of the bouts sampled. Methods to estimate the 'true' repertoire size are unreliable when singletons exceed doubletons. Birds with >20 bouts have been excluded as outliers from B and C.

Table 1. **Definitions of behavioral elements.** See also Figure 1 and Myers et al. (2019).

Behavioral element	letter	Description
Short descent	*a*	Descent of 2-3m in a shallow half-U shape.
Short ascent	*b*	Ascent of 2-3m in a shallow half-U shape. Usually performs writhing at the same time.
Writhing	*c*	While performing a short ascent (*b*), the bird shudders up and down while flipping its tail, and makes modified pulses of wing sound called 'chirrups'.
Long ascent	*d*	During a dive display, an ascent of about 20m.
Medium ascent	*d'*	Ascent in between *b* and *d* in height. Hybrids only (the 'mini-dive' display type)
Long descent	*e*	During a dive, after element g, male descends, traces similar path as ascent, spreads tail feathers near end of descent, produces dive sound; finishes descent.
Shuttle segment	*f*	In a hover or near hover, bird flies in a bouncy, repeating motion and makes one set of modified pulses of wing sound (chirrups). Occurs within a shuttle display.
Turn-around	*g*	Male turns around roughly 180 degrees, so that a subsequent element is in the opposite direction of the previous movement.

Table 2. **Number of birds expressing each display type,** including typical and rare variants of displays (24 total). Also see Figure 3. Raw number of displays inside parentheses.

Display	Letter	Element sequence	Rufous	Hybrid	Allen's
Dummy (sequence start/end)	E	N/A	N/A	N/A	N/A
SHUTTLE	S	*fg**	44 (226)	301 (1,961)	34 (160)
HALF PENDULUM (typical)	H	*bcg*	39 (139)	285 (1,410)	30 (91)
Half pendulum, no writhing	H_X	*b_g*	0	3 (3)	0
Half pendulum, double writhing	H_W	*bcbcg*	1 (1)	2 (2)	0
Half pendulum, no turn-around	H_D	*bc_*	0	4 (4)	0
RUFOUS DIVE (typical)	R	*adgebcg*	46 (840)	215 (2,468)	4 (10)
Rufous dive, no writhing at end	R_X	*adgeb_g*	6 (10)	33 (42)	6 (9)
Rufous dive, aborted	R_A	*adg____*	14 (28)	48 (73)	0
Rufous dive, mini	R_M	*ad'gebcg*	0	3 (29)**	0
Rufous dive, no short descent	R_L	*_dgebcg*	6 (8)	11 (11)	0
Rufous dive, no short descent, aborted	R_{LA}	*_dg____*	2 (2)	0	0
ALLEN'S DIVE (typical)	A	*abcdgebcg*	0	194 (1,308)	32 (165)
Allen’s dive, no end writhing	A_X	*abcdgeb_g*	0	37 (78)	18 (115)
Allen’s dive, aborted	A_A	*abcdg____*	0	44 (73)	7 (19)
Allen’s dive, mini***	A_M	*abcd'gebcg*	0	3 (4)	0
Allen’s dive, no short descent	A_L	*_bcdgebcg*	0	11 (17)	1 (1)
Allen's dive, no short descent & aborted	A_{LA}	*_bcdg____*	0	3 (3)	0
Allen's dive, no short descent & no end writhing	A_{LX}	*_bcdgeb_g*	0	2 (2)	0
Allen's dive, no turn-around	A_D	*abcdgebc_*	0	1 (1)	0
PENDULUM (typical)	P	*abcg*	1 (1)	236 (8,915)	35 (2,562)
Pendulum, no writhing	P_X	*ab_g*	0	43 (87)	2 (2)
Pendulum, no turn-around	P_D	*abc_*	0	8 (16)	0
Pendulum, double writhing	P_W	*abcbcg*	0	4 (4)	1 (1)
TOTAL: birds (displays)			46 (1,256)	306 (16,510)	35 (3,135)

* shuttle display element sequence not used in this study

** The mini-dive: element d' is intermediate between *b* (short ascent) and *d* (long ascent)

*** An outlier, E08526, performed 21 R_M

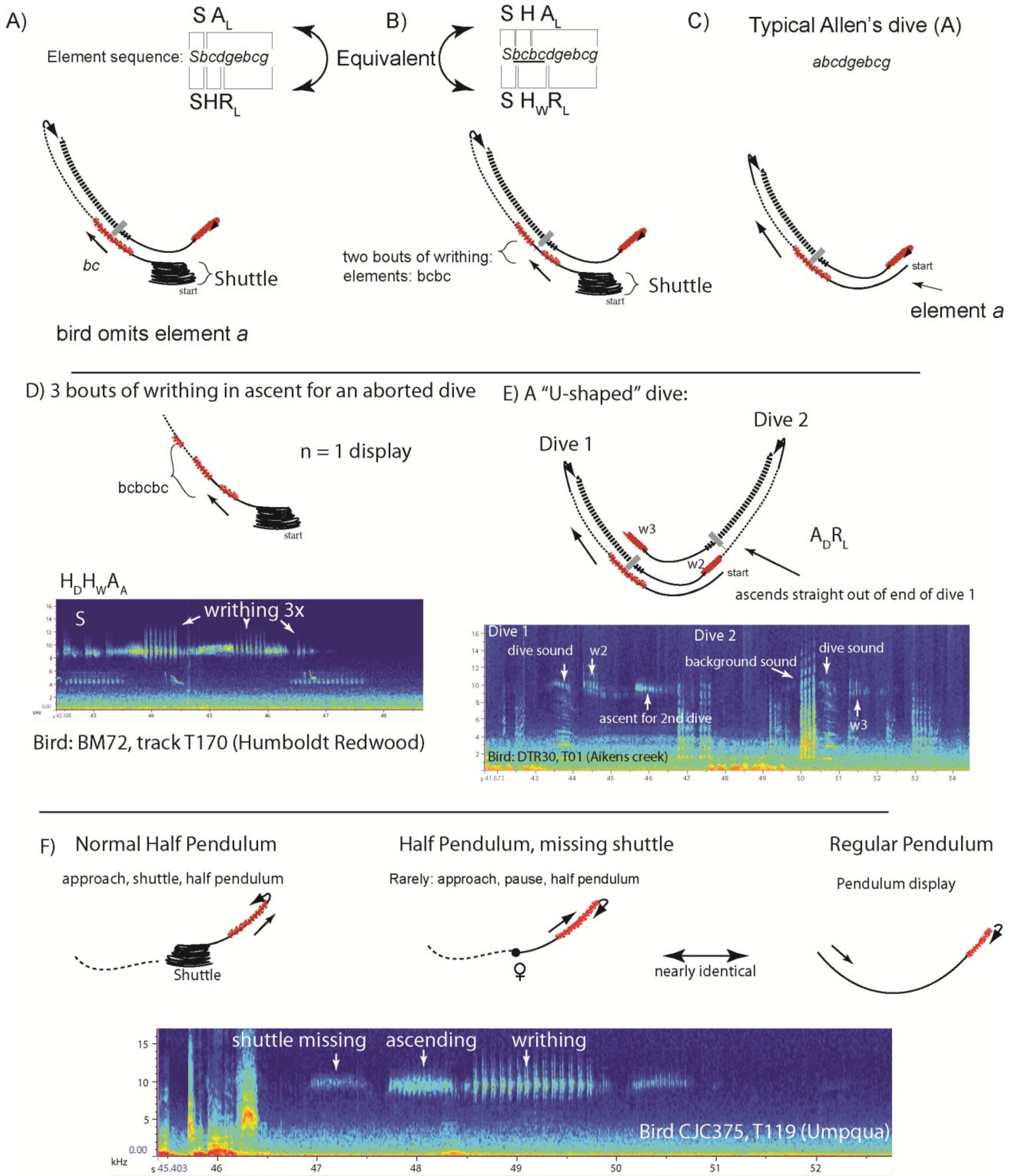

Figure S1 Ambiguous or noteworthy display sequences. A) The element sequence *Sbcdgebcg* (a bird ascends for a dive straight out of a shuttle display; element *a* is missing; compare with normal dive shown in C) could be equivalently interpreted as SA_L or SHR_L. **B**) The element sequence *Sbcbcdgebcg*, in which a bird did two bouts of writhing (bcbc) in the ascent, could be interpreted as SHA_L or SH_WR_L. Since H_W and R_L are both rare variants, while SH is common and A_L is rare, it is more parsimonious to code this sequence as SHA_L. Therefore, we coded the sequence in A) as SA_L.

D) Three bouts of writhing in an ascent for an aborted dive. **E**) U-shaped dive. Dive sound (in spectrogram) is made by tail-feathers. Background sound are calls from other birds. **F**) Half pendulums normally follow a shuttle display. Rarely, the male would approach the caged female (dotted line; in spectrogram, "shuttle missing"), but then launch into a half pendulum. Thus, display 3-mers such as "EEH" were possible. Acoustically, this sequence sounds like a pendulum display; thus display sequences scored only by sound (not visually observed) will tend to be mis-scored as pendulum displays. Therefore, display sequences in which half pendulums did not come after a shuttle are likely under-represented in our dataset.

Table S1. List of 3-mers expressed by at least one bird. See text for impossible 3-mers (e.g. EEE or SS* are not possible) (# of birds expressing the 3-mer are in parentheses).

3-mer	Allen's	Hybrids	Rufous	Notes
AAA	0 (0)	54 (34)	0 (0)	
AAE	11 (4)	188 (85)	0 (0)	
AAP	2 (2)	15 (12)	0 (0)	
AAR	0 (0)	1 (1)	0 (0)	
AAS	0 (0)	10 (6)	0 (0)	
AEE	267 (35)	992 (192)	0 (0)	
APA	5 (5)	30 (24)	0 (0)	
APE	3 (3)	39 (30)	0 (0)	
APP	10 (6)	63 (45)	0 (0)	
APR	0 (0)	9 (9)	0 (0)	
APS	0 (0)	16 (12)	0 (0)	
ARA	1 (1)	0 (0)	0 (0)	
ARE	0 (0)	13 (12)	0 (0)	
ARR	0 (0)	1 (1)	0 (0)	
ARS	0 (0)	2 (2)	0 (0)	
ASA	0 (0)	1 (1)	0 (0)	
ASE	0 (0)	12 (12)	0 (0)	
ASH	1 (1)	40 (30)	0 (0)	
EAA	0 (0)	14 (12)	0 (0)	
EAE	2 (2)	18 (10)	0 (0)	
EAP	0 (0)	7 (6)	0 (0)	
EAR	0 (0)	2 (2)	0 (0)	
EAS	0 (0)	4 (4)	0 (0)	
EEA	2 (2)	45 (24)	0 (0)	
EEH	0 (0)	5 (3)	1 (1)	Unusual for H to not be preceded by S (see text)
EEP	396 (35)	1656 (215)	1 (1)	Presence in Rufous
EER	3 (2)	704 (129)	305 (46)	
EES	95 (28)	1013 (257)	173 (40)	
EHE	0 (0)	1 (1)	1 (1)	Unusual for H to not be preceded by S (see text)
EHP	0 (0)	1 (1)	0 (0)	Unusual for H to not be preceded by S (see text)
EPA	5 (4)	83 (49)	0 (0)	

EPE	18 (11)	58 (42)	1 (1)	
EPP	363 (35)	1300 (200)	0 (0)	
EPR	1 (1)	153 (73)	0 (0)	
EPS	9 (9)	65 (46)	0 (0)	
ERA	0 (0)	6 (4)	0 (0)	
ERE	3 (2)	78 (60)	51 (30)	
ERP	0 (0)	41 (19)	1 (1)	
ERR	0 (0)	510 (103)	246 (46)	
ERS	0 (0)	65 (44)	7 (6)	
ESA	0 (0)	8 (4)	0 (0)	
ESE	47 (22)	341 (142)	69 (26)	Entire bout of display is a shuttle display
ESH	45 (23)	642 (220)	98 (35)	
ESP	3 (2)	8 (7)	0 (0)	Unusual for no H between S and P
ESR	0 (0)	18 (17)	5 (5)	there is no H after S; 3-mer type *SR only possible when R is type R_L
HAA	0 (0)	6 (5)	0 (0)	
HAE	3 (3)	54 (41)	0 (0)	
HAP	0 (0)	3 (3)	0 (0)	
HAR	0 (0)	2 (2)	0 (0)	
HEE	23 (13)	371 (176)	56 (27)	
HHE	0 (0)	2 (2)	0 (0)	Unusual for H to follow H (see text)
HHR	0 (0)	2 (2)	0 (0)	Unusual for H to follow H (see text)
HHS	0 (0)	2 (2)	0 (0)	Unusual for H to follow H (see text)
HPA	6 (4)	83 (56)	0 (0)	
HPE	14 (10)	137 (93)	0 (0)	
HPH	0 (0)	2 (2)	0 (0)	
HPP	37 (20)	287 (125)	0 (0)	
HPR	1 (1)	35 (24)	0 (0)	
HPS	2 (2)	21 (15)	0 (0)	
HRA	0 (0)	2 (2)	0 (0)	
HRE	0 (0)	104 (62)	40 (20)	
HRH	0 (0)	2 (2)	0 (0)	
HRP	0 (0)	7 (7)	0 (0)	
HRR	0 (0)	68 (40)	13 (9)	

HRS	0 (0)	5 (5)	2 (2)	
HSA	0 (0)	2 (2)	0 (0)	no H after S; 3-mer type *SA only possible when A is type A_L
HSE	1 (1)	34 (25)	5 (5)	
HSH	4 (3)	186 (86)	21 (12)	
HSR	0 (0)	1 (1)	3 (3)	there is no H after S; 3-mer type *SR only possible when R is type R_L
PAA	13 (4)	182 (82)	0 (0)	
PAE	249 (35)	681 (172)	0 (0)	
PAP	15 (12)	126 (75)	0 (0)	
PAR	1 (1)	10 (10)	0 (0)	
PAS	1 (1)	41 (34)	0 (0)	
PEE	128 (33)	581 (195)	1 (1)	
PHE	0 (0)	3 (2)	0 (0)	
PPA	263 (35)	836 (174)	0 (0)	
PPE	92 (29)	308 (138)	0 (0)	
PPP	1671 (35)	4715 (198)	0 (0)	
PPR	12 (7)	225 (94)	0 (0)	
PPS	47 (16)	350 (133)	0 (0)	
PRA	1 (1)	45 (36)	0 (0)	
PRE	9 (7)	235 (98)	0 (0)	
PRH	0 (0)	1 (1)	0 (0)	
PRP	3 (2)	114 (52)	0 (0)	
PRR	1 (1)	53 (28)	1 (1)	
PRS	0 (0)	40 (28)	0 (0)	
PSA	1 (1)	5 (5)	0 (0)	
PSE	17 (8)	67 (48)	0 (0)	
PSH	41 (16)	400 (137)	0 (0)	
PSP	0 (0)	1 (1)	0 (0)	
PSR	0 (0)	2 (2)	0 (0)	
RAA	0 (0)	11 (8)	0 (0)	
RAE	1 (1)	36 (33)	0 (0)	
RAP	1 (1)	6 (6)	0 (0)	
RAR	0 (0)	1 (1)	0 (0)	

RAS	0 (0)	3 (3)	0 (0)	
REE	13 (8)	974 (205)	344 (46)	
RHE	0 (0)	3 (3)	0 (0)	
RPA	0 (0)	5 (5)	0 (0)	
RPE	1 (1)	41 (29)	0 (0)	
RPP	1 (1)	61 (36)	0 (0)	
RPR	0 (0)	65 (31)	1 (1)	
RPS	1 (1)	24 (13)	0 (0)	
RRA	0 (0)	3 (3)	0 (0)	
RRE	1 (1)	528 (117)	248 (45)	
RRP	0 (0)	32 (19)	0 (0)	
RRR	0 (0)	556 (87)	256 (41)	
RRS	0 (0)	80 (44)	15 (10)	
RSA	0 (0)	2 (2)	0 (0)	there is no H after S; 3-mer type *SA only possible when A is type A_L
RSE	0 (0)	44 (32)	4 (4)	
RSH	0 (0)	137 (68)	20 (11)	
RSR	0 (0)	9 (7)	0 (0)	there is no H after S; 3-mer type *SR only possible when R is type R_L
SAA	0 (0)	1 (1)	0 (0)	there is no H after S; 3-mer type *SA only possible when A is type A_L
SAE	1 (1)	17 (12)	0 (0)	
SEE	65 (23)	496 (194)	78 (29)	
SHA	3 (3)	65 (43)	0 (0)	
SHE	23 (13)	364 (172)	55 (26)	
SHH	0 (0)	7 (6)	0 (0)	Unusual for H to follow H (see text)
SHP	60 (26)	563 (178)	0 (0)	
SHR	0 (0)	186 (84)	55 (25)	
SHS	5 (3)	222 (96)	29 (15)	
SPE	0 (0)	1 (1)	0 (0)	
SPP	3 (2)	6 (5)	0 (0)	
SPR	0 (0)	1 (1)	0 (0)	
SPS	0 (0)	1 (1)	0 (0)	
SRE	0 (0)	16 (13)	5 (5)	there is no H after S; 3-mer type SR* only

				possible when R is type R_L
SRP	0 (0)	2 (2)	0 (0)	there is no H after S; 3-mer type SR* only possible when R is type R_L
SRR	0 (0)	12 (12)	3 (3)	there is no H after S; 3-mer type SR* only possible when R is type R_L