Title: Spatial variation in population genomic responses to over a century of anthropogenic change within a tidal marsh songbird

Authors: Phred M. Benham[1], Jennifer Walsh[2], Rauri C. K. Bowie[1]

Author affiliations:

(1) Museum of Vertebrate Zoology & Department of Integrative Biology, University of California, Berkeley, Berkeley, CA

(2) Fuller Evolutionary Biology Program, Cornell Lab of Ornithology, Cornell University, Ithaca, NY.

ABSTRACT: Human activities have wide-ranging impacts on natural populations, but our ability to pinpoint population responses to these activities is often limited to the analysis of populations studied well after the fact. Museum collections can provide critical baseline data on patterns of genetic diversity, connectivity, and population structure prior to the onset of human perturbation. Here, we leverage a spatially-replicated time series of specimens to document population genomic responses to the destruction of ~90% of coastal habitats occupied by the Savannah sparrow (*Passerculus sandwichensis*) in California. We sequenced 219 sparrows collected from 1889-2017 across the state of California using an exome capture approach. Spatial-temporal analyses of genetic diversity indicate that sparrow populations from southern California historically exhibited lower levels of genetic diversity and have also experienced the most significant temporal declines in genetic diversity. In contrast, birds from the San Francisco Bay area exhibit near constant levels of genetic diversity, despite experiencing similar levels of habitat loss. Stable genetic diversity in

Bay Area birds may stem from increased levels of gene flow from interior California populations. While greater levels of gene flow may prevent declines in genetic diversity, we also find that immigration from inland freshwater-adapted populations into tidal marsh populations has led to the erosion of divergence in loci associated with tidal marsh adaptation. Although gene flow can play an important role in mitigating inbreeding depression, our results point to the potential negative consequences of gene flow across environmental gradients on locally adapted populations.

KEYWORDS: conservation genomics; tidal marshes; Passerellidae; anthropogenic change; museum collections, exome capture

SIGNIFICANCE: Museum collections preserve a record of population responses to anthropogenic change. We analyzed genomic data from historic specimens to document the responses of coastal California Savannah sparrows to rampant tidal marsh loss over the past century. We find that temporal shifts in source-sink dynamics among populations allowed some populations to maintain high genetic diversity in the face of habitat loss, yet this led to the erosion of genetic divergence at loci associated with tidal marsh adaptation. Shifting patterns of gene flow through time in response to habitat loss may result in negative fitness consequences and outbreeding depression. This has important implications for the long-standing debate on assisted gene flow in populations threatened with inbreeding depression.

INTRODUCTION: The dynamics of local adaptation, gene flow, and effective population size influence how species respond to anthropogenic change in complex ways. Fragmentation and

habitat degradation can lead to population declines, increased isolation, and increased risk of inbreeding depression (e.g., Tallmon et al. 2004; Frankham 2005; Whitely et al. 2015). On the other hand, changing patterns of gene flow (i.e. directions, rate) among populations locally adapted to different environments may lead to the influx of mal-adapted immigrants and outbreeding depression (e.g. Edmands et al. 2007; Frankham et al. 2011). It is also possible that increased gene flow, even from a maladapted source population, could contribute to the evolutionary rescue of a peripheral small population by providing the essential allelic variation needed to facilitate adaptation to environmental change (e.g. Bell et al. 2019). Mixed support exists for all these outcomes, but a deeper understanding of the ecological and demographic context in which certain processes, such as gene flow, may be beneficial or detrimental to threatened populations is sorely needed.

It can take several generations to observe how these different scenarios might shape population responses to human-mediated ecological change. For many long-lived, non-model, and threatened taxa facing immediate conservation threats, we do not have the luxury of waiting to see if our current models of population responses will prove correct. Time-series data spanning past human transformation of the landscape can inform this issue through documenting how natural populations have been impacted by past episodes of ecological change. Museum collections are important repositories of time-series data for evaluating population responses to anthropogenic change across broad spatial, temporal, and taxonomic scales (e.g. Holmes et al. 2016; Billerman & Walsh 2019). Such collections provide the opportunity for analyses across spatially-replicated, time-series that can complement insights from other sources of time-series data (e.g. long-term ecological studies; Magurran et al. 2010). The perspective provided by time-series analyses will

68 ultimately contribute to more accurate predictions of population responses to future anthropogenic
69 change.

70 Historical DNA isolated from museum specimens has proven to be a critical resource for
71 establishing baseline levels of genetic diversity and population structure (Wanderler et al. 2007).
72 More recent analyses have increasingly leveraged genomic datasets with dramatically increased
73 power to address a host of conservation genetics questions (Bi et al. 2013; Suchan et al. 2016; Bi
74 et al. 2019; Feng et al. 2019; van der Valk et al. 2019; Sánchez-Barreiro et al. 2021). In particular,
75 studies comparing patterns of temporal change at both neutral loci and loci potentially under
76 selection can provide further insights into: (1) how populations adapt to human-mediated
77 ecological change; and (2) how changing patterns of gene flow, population size, and habitat
78 degradation influence variation at neutral versus functional loci. Here, we leverage spatially
79 replicated time-series of Savannah sparrow (*Passerculus sandwichensis*) specimens collected from
80 throughout the state of California, to analyze genomic responses across a landscape of spatially
81 heterogenous anthropogenic change.

82 Savannah sparrows are one of the most widespread North American songbirds, with 17
83 subspecies breeding across a range of open habitats from Alaskan tundra to the highlands of
84 Guatemala (Fig. 1a; Wheelwright & Rising 2008). Within California, there are four subspecies
85 (Grinnell & Miller 1944; Fig. 1a): *Passerculus sandwichensis brooksi* and *P. s. nevadensis* are
86 both migratory subspecies that occupy grasslands in extreme northwest and eastern California,
87 respectively. *P. s. alaudinus* resides in tidal marsh and coastal grassland habitats from Humboldt
88 Bay (Humboldt County) south to Point Conception (Santa Barbara County); and *P. s. beldingi* is
89 a year-round resident of tidal marshes from Point Conception south to Baja California. These
90 coastal populations exhibit local adaptation to high salinity and flooding in tidal marshes with prior

work documenting physiological, behavioral, and genomic divergence from other inland, freshwater-associated California populations of the species (Cade & Bartholomew 1959; Poulson & Bartholomew 1962; Walsh et al. 2019; Benham & Cheviron 2020; Benham & Bowie 2021).

California Savannah sparrows span a landscape that has experienced dramatic but spatially variable transformations due to human activity. The human population of California has grown by more than 38 million since 1900 (U.S. Census Bureau 2019). Much of this growth is concentrated in coastal and central California, while many areas of eastern and northern California retain extensive tracts of wilderness. Similarly, maximum temperatures increased the most in urban and agricultural areas over the past century, but decreased along the immediate coast (e.g., LaDochy et al. 2007; Lebassi et al. 2008). Coastal populations of the Savannah sparrow have been particularly impacted by anthropogenic change, with agriculture, urbanization, and certain industries (e.g., salt evaporation mining) collectively destroying 75% of the state's coastal wetlands between the late 1800s and the present (Powell 2006). This includes critical tidal marsh habitat for many Savannah sparrow populations, with the loss of tidal marsh habitat in some estuaries exceeding 90% (e.g., San Francisco Bay; Marshall & Dedrick 1994). Consequently, coastal Savannah sparrow species are listed as either state endangered species (*P. s. beldingi;* Zembal & Hoffmann 2010) or bird species of special concern (*P. s. alaudinus;* Fitton 2008).

To explore temporal patterns of genetic change across these different Savannah sparrow populations, we sequenced genomic data from 219 individuals sampled over a 127 year period, representing nine spatially replicated time-series. We analyzed this dataset to address the following questions: (1) how has tidal marsh habitat loss impacted patterns of population structure and gene flow among California Savannah sparrow populations? (2) How has genetic diversity changed through time across different Savannah sparrow populations? And (3) what impact has changing

114 population demography had on loci associated with adaptation to tidal marsh habitats? By tackling
115 these questions we aim to document the nuances of past population responses to anthropogenic
116 change. This will contribute to an improved understanding of the demographic and ecological
117 factors that predict population resilience or decline in the face of ongoing and future environmental
118 change.
119
120 **RESULTS:**
121 *Patterns of population structure:* We extracted DNA and prepared libraries for sequencing with a
122 custom designed exome capture array (see supplemental methods). DNA was extracted from the
123 toepads of 155 historic specimens and blood and tissues from 94 modern samples (74 from
124 California and 20 from outside; Dataset S1). Captures performed well with an average of 73.4%
125 (modern) and 74.8% (historic) of cleaned reads mapping to the custom target capture regions
126 (supplemental Table S1, Figure S1). Average coverage was five times higher in modern samples
127 (range: 40.5-73.7; mean: 55.0x) than historical samples (range: 2.0-29.4; mean: 11.0x). To account
128 for this discrepancy in coverage for some downstream analyses, we randomly downsampled reads
129 of modern individuals to generate a dataset where modern birds had an average coverage of 10x
130 (supplemental Figure S2). Unlike most ancient DNA datasets, C to T and G to A mutations were
131 slightly more common in the modern dataset than historic dataset (supplemental Figure S3), which
132 likely reflects the influence of the USER enzyme treatment on historic samples.
133 Principal components analysis of the full dataset (60,765 SNPs) identified five genetic
134 clusters. Modern and historical samples collected from the same locality clustered together (Fig.
135 1b). The first principal component explained 2.8% of the variance in the data and separated out
136 three clusters: (1) *Passerculus sandwichensis beldingi* subspecies from south coastal California;

137 (2) birds from Morro Bay in San Luis Obispo County; and (3) birds from the rest of California.
138 The second principal component explained 1.1% of the variance in the data and divided birds from
139 the *P. s. nevadensis* subspecies in eastern California, from birds in northwest California (*P. s.*
140 *alaudinus* and *P. s. brooksi*), and the remaining birds in the *P. s. alaudinus* subspecies from the
141 Bay Area and central California coast. The third principal component (0.9% of variance) separated
142 modern from historical samples and the fourth component (0.8% of variance) split northwest
143 California birds from other populations. PCA of other datasets (downsampled coverage, CpG sites
144 removed, C->T and G->A sites removed) revealed similar patterns of population clustering
145 (supplemental Figure S4).

146 To test for temporal changes in population structure and admixture among California
147 Savannah sparrow populations, we analyzed a dataset of 65,583 unlinked SNPs and 239
148 individuals (20 from outside California) within the program DyStruct (Joseph & Pe'er 2019).
149 DyStruct resembles other model-based clustering programs (e.g. Alexander et al. 2009), but
150 explicitly accounts for the temporal dynamics of allele frequency change in time-series data. After
151 running DyStruct three times independently for K equals 1 to 12 population clusters, we found
152 highly similar likelihoods for K=3 to 5 (log-likelihood= -1383.63 to -1383.76; Fig. 1c). All three
153 values of K identified *P. s. beldingi* from south coastal California and Mexico as a distinct cluster,
154 *P. s. alaudinus* of the central and northern California coast was a second genetic cluster, with *P. s.*
155 *nevadensis* of eastern California together with birds from Wyoming, Alaska, and Washington state
156 forming a third cluster. The main difference between K=3 and 4 was that Morro Bay birds were
157 recognized as a distinct cluster, and at K=5 modern birds from the central and northern California
158 coast were identified as a distinct cluster. This analysis also revealed temporal variation in
159 population structure with modern birds from the central and northern California coast exhibiting a

160 17% (Morro Bay) to 64% (Humboldt Bay) increase in the proportion of ancestry shared with inland
161 eastern California birds. Temporal changes in ancestry proportions points to a recent increase in
162 the level of gene flow from inland eastern California populations into coastal populations.
163
164 *Spatial and temporal patterns of genetic diversity and gene flow:* Visualization of effective
165 migration and diversity surfaces output from the program EEMS (Petkova et al. 2016) show similar
166 patterns of migration and genetic diversity among California Savannah sparrow populations in
167 both the historical (<1940) and modern (>2000) datasets (Fig. 2a; supplemental Fig. S5). In both
168 historic and modern datasets, southern California birds exhibited lower effective diversity rates
169 (the expected genetic dissimilarity between two individuals in a population) and greater isolation
170 relative to northern California birds.
171 We estimated the deviation from expected heterozygosity, F, as a measure of inbreeding
172 for both exons (n=9,971) and nongenic (n=1,024) regions. These estimates indicated that southern
173 populations at Newport and Morro Bays were more inbred than northern populations (Fig. 2b). In
174 all populations, we also detected a significant decrease in inbreeding over the past century,
175 suggesting that these populations have become more outbred. In the Newport Bay population (*P.*
176 *s. beldingi*), the inbreeding coefficient significantly decreased from the early 1900s to the 1960s
177 and then again from the 1960s to the 2010s. This suggests that the state endangered *P. s. beldingi*
178 has become more outbred despite increased habitat fragmentation over the past century.
179 Newport and Morro Bay populations also exhibited lower levels of genome-wide
180 nucleotide diversity (π), Watterson's theta (θ) and higher Tajima's D than populations from
181 northern California (Fig. 2c). These two populations exhibited significant declines in π and θ over
182 time within exons. Newport Bay also saw a significant increase in Tajima's D. Patterns were

183 similar for nongenic data, but these changes were not statistically significant. Despite experiencing
184 comparable levels of tidal marsh loss, sparrows from the San Francisco Bay Area exhibited no
185 significant change in π for exon or non-genic data but did experience a significant decline in θ for
186 both exons and non-genic data. However, despite these declines in θ, Bay Area birds retain
187 comparable levels of π and θ to birds from NW and eastern California. The NW and eastern
188 California populations exhibited either no change in these metrics or slight increases (e.g.,
189 temporal increases in π in NW California). Estimates of these parameters using datasets that were
190 less conservatively filtered showed similar patterns (supplemental Fig. S5). For Newport Bay
191 birds, significant decreases in π and θ only occurred between the 1968 (1960s) and 2016 time
192 points (supplemental Fig. S6). San Francisco Bay on the other hand showed a long-term increase
193 in Tajima's D from the 1900s-1910s to the present, which appeared to be driven by decreasing θ
194 through time (supplemental Fig. S7).

195 We next fit an isolation with migration model to the joint site frequency spectrum of
196 Newport Bay, eastern California, and the Bay Area from both the historical (pre-1920) and modern
197 (post-2000) periods in *Moments* (Jouganous et al. 2017). Historical and modern estimates of
198 divergence time among the three populations, effective population sizes, and gene flow patterns
199 were broadly overlapping (supplemental Fig. S8). Divergence between *P. s. beldingi* and northern
200 California occurred about 340,000 years ago and divergence between the Bay Area and inland
201 eastern California ~35,000 years ago. Estimates of effective population size for both time periods
202 was ~10,000 for *P. s. beldingi,* 61-68,000 for *P. s. alaudinus,* and 85-108,000 for *P. s. nevadensis.*
203 Both historic and modern gene flow rates into Newport Bay from both eastern California and the
204 Bay Area was less than 0.5 migrants per generation (supplemental Fig. S8). In contrast, gene flow
205 between the northern California populations was much higher and asymmetric with gene flow

206 from the Bay Area to eastern California exceeding 12 migrants per generation, but only 1-3
207 migrants per generation in the reverse direction.
208
209 *Selection analyses:* Mean Fst divergence between each of the four historical tidal marsh
210 populations (Humboldt Bay, Bay Area, Morro Bay, Newport Bay) and inland eastern California
211 populations varied from 0.014 at Humboldt Bay to 0.098 at Newport Bay in southern California
212 (Table 1). Modern values of Fst divergence showed increases in Morro Bay from 0.054 to 0.062,
213 but a sharp decrease in mean Fst at Humboldt Bay from 0.014 to -0.002. Across the four tidal
214 marsh populations we identified between 203 (Newport Bay) and 229 SNPs (Bay Area) that
215 exceeded the 99th percentile threshold (Fig. 3a). In all four tidal marsh populations we detected a
216 significant temporal decrease in Fst divergence for outlier SNPs relative to non-outlier SNPs (Fig.
217 3b). Analyses based on loci that were Fst, Dxy, and LFMM outliers showed similar patterns
218 (supplemental Fig. S9). No genes were found to be outliers in all four coastal populations, but two
219 genes were outliers in three populations (supplemental Table S2). This includes the estrogen
220 receptor 1 gene (ESR1), which was an outlier in all populations except Humboldt Bay. Moreover,
221 a T->C mutation in the 3' UTR region in the final exon of ESR1 exhibited the greatest Fst
222 divergence in both Newport Bay (Fig. 3c; Fst=0.73) and the Bay Area (Fig. 3d; Fst=0.41). While
223 this SNP showed large increases in Fst divergence within Morro Bay (Fst_{hist}: 0.45; Fst_{mod}: 0.78;
224 ΔFst: 0.33), it exhibited large temporal declines in the Bay Area (ΔFst=-0.17). Temporal change
225 in Fst at this locus is likely due to decreasing frequencies of the alternate C allele of this SNP in
226 the Bay Area (change in frequency -0.24; Fig. 3e).
227 Increased gene flow from inland eastern California populations into northern coastal
228 populations should also lead to the increased homogenization of allele frequencies among

229 populations towards the present. Further, if outlier loci associated with local adaptation to tidal
230 marshes exhibit increased resilience to gene flow then we would expect less homogenization of
231 allele frequencies at these loci relative to non-outlier loci. We tested these predictions by
232 estimating the degree of correlation between historical deviation in allele frequency and change in
233 allele frequency through time for each SNP (Gompert et al. 2021). In this analysis, stronger
234 correlations will reflect greater levels of gene flow. For each of the four tidal marsh populations
235 we estimated deviation from mean allele frequency of the tidal marsh and eastern California
236 populations as $\overline{p_i} - p_{ij.}$ Where $\overline{p_i}$ is the mean allele frequency of the ith SNP between inland eastern
237 California and the jth tidal marsh population and p_{ij} is the allele frequency of the ith SNP in the jth
238 tidal marsh population. Difference in allele frequency change through time was estimated for each
239 SNP in each tidal marsh population as $p_{ij}^{modern} - p_{ij}^{historical}$. We used permutation tests to test for
240 significant differences from zero (no correlation) and between outlier and non-outlier SNPs. This
241 showed that correlations between historic deviation and temporal change in allele frequency were
242 significantly different from zero for both the Bay Area (mean correlation: 0.17; p-value=0.015)
243 and Humboldt Bay tidal marsh populations (mean: 0.30; p-value=0.001), but not significant for
244 Morro Bay (mean: 0.16; p-value=0.08) or Newport Bay (mean: 0.05; p-value: 0.33; Fig. 4a).
245 Second, we found that outlier loci exhibit significantly more positive correlations relative to non-
246 outlier loci across all four coastal populations (Fig. 4a). Finally, the degree of correlation varies
247 significantly with the extent to which inland eastern California ancestry increased in each tidal
248 marsh population through time (r^2=0.92; p-value=0.03; Fig. 4b). This result remained regardless
249 of the threshold used to identify outlier SNPs (thresholds: 0.9, 0.95, 0.99th percentile of Fst
250 divergence). In sum, ΔFst and temporal allele frequency change results both indicate that the
251 observed increase in gene flow levels from the freshwater-adapted, eastern California population

252 into tidal marsh populations eroded divergence at loci potentially associated with adaptation to a
253 high salinity environment.
254
255 **DISCUSSION:**

256 There is a pressing need to understand the ecological and demographic factors that
257 influence population genetic responses to habitat loss and environmental change. Through the
258 preservation of baseline genetic diversity data from before the onset of large-scale human
259 disturbance, museum collections are poised to play an increasingly important role in addressing
260 this issue. We leveraged replicated, time-series of specimens to document the impacts of habitat
261 loss in tidal marsh populations of the Savannah sparrow in California. Over 75% of coastal
262 wetlands in California have been lost due to human activities (Powell 2006), but the magnitude of
263 tidal marsh loss varies extensively among estuaries from 90% loss of marshes in the San Francisco
264 Bay (Marshall & Dedrick et al. 1994; Takekawa et al. 2006) to 84% increases of marsh habitat
265 over the past century in Morro Bay (Gerdes et al. 1974). This variation in habitat loss was not
266 found to predict declines in genetic diversity across different Savannah sparrow populations.
267 Rather, we show that both southern tidal marsh populations at Morro and Newport Bays: (1)
268 historically exhibited lower levels of genetic diversity; (2) higher inbreeding coefficients; and (3)
269 greater temporal declines in genetic diversity relative to northern California populations (Fig. 2).
270 In contrast, despite experiencing 90% tidal marsh habitat loss, San Francisco Bay sparrows still
271 exhibit similar levels of genetic diversity to other northern California populations (Fig. 2c; Supp.
272 Fig. S7). While habitat loss and fragmentation likely played some role in declining genetic
273 diversity, our analyses point to more complex demographic dynamics contributing to patterns of
274 changing genetic diversity among these populations.

275 In other species, a disconnect between habitat loss and genetic diversity declines has been
276 attributed to a range of different demographic, life history, ecological, and behavioral factors (e.g.
277 Welch et al. 2012; Lino et al. 2019). Many of these factors are unlikely to contribute to the
278 observed differences in genetic diversity between southern and northern California tidal marshes
279 as generation time, body size, and migratory behavior are all similar across coastal California
280 Savannah sparrow populations. Differences in the degree of specialization to tidal marshes could
281 contribute to differences in genetic diversity declines. Tidal marsh habitats are the stronghold for
282 *P. s. alaudinus* populations in northern California, but this subspecies can regularly be found
283 breeding in coastal grasslands (Grinnell & Miller 1944). This greater niche breadth may buffer
284 them from tidal marsh habitat losses. Indeed, populations from Humboldt Bay shifted from tidal
285 marsh into pasture habitats as pasture land expanded at the expense of tidal marsh habitat (Fitton
286 2008). In addition to differences in habitat use, our demographic and population structure analyses
287 indicate that spatial-temporal patterns of gene flow may have contributed to differences in genetic
288 diversity declines observed among coastal Savannah sparrow populations.

289 EEMS and *Moments* analyses suggest that more gene flow occurs among northern
290 California populations than between northern and southern California (supplemental Fig. S5, S9).
291 The greater isolation of Newport Bay birds may play a role in the greater declines of genetic
292 diversity observed in southern California despite similar levels of historic habitat loss. EEMS and
293 *Moments* did not show temporal differences in migration rates among populations. In contrast,
294 DyStruct analyses point to a reversal in the direction of gene flow from historically greater
295 immigration into eastern California from the coast to greater gene flow rates from eastern into
296 coastal California populations toward the present. These contrasting results likely stem from the
297 isolation with continuous migration model tested in *Moments,* which would return average

298 migration rates since population divergence and not necessarily identify recent shifts in gene flow
299 rates. Further, the more dynamic history of shifting gene flow patterns identified with DyStruct
300 was supported by declines in mean genome wide Fst divergence between Humboldt Bay and
301 eastern California (from 0.014 to 0) and between the Bay Area and eastern California (from 0.015
302 to 0.013). Second, allele frequency analyses showed greater positive correlations between
303 historical deviations from mean allele frequency and temporal change in allele frequency existed
304 for the Bay Area and Humboldt Bay (Fig. 4). Morro Bay and Newport Bay in southern California
305 were not significantly different from 0 (no correlation). These patterns were consistent with the
306 expectation that greater gene flow leads to greater correlation between the direction and magnitude
307 of allele frequency change (Gompert et al. 2021).

308 Although habitat loss is typically thought to increase population fragmentation and reduce
309 migration, theoretical work suggests that habitat loss and subsequent population declines can also
310 leave populations more susceptible to hybridization and gene flow (e.g. Todesco et al. 2016).
311 Differences in effective population sizes between populations will influence the direction of gene
312 flow with theoretical and empirical evidence suggesting that the direction of gene flow will
313 typically be from larger into smaller populations (Currat et al. 2008). Our demographic analyses
314 indicate that historically ~3 migrants per generation immigrated from inland eastern California
315 into the Bay Area and 13.5 migrants per generation immigrated in the opposite direction. If habitat
316 loss led to declining local recruitment and population growth within the Bay Area, the impact of
317 ongoing gene flow from inland eastern California into the Bay Area would become more
318 significant and could explain the observed increase in eastern California ancestry. These patterns
319 suggest that source-sink dynamics have shifted through time with the Bay Area potentially

320 becoming a sink population at present, sustained by migration from eastern California (Dias et al.
321 1996; Gaggiotti & Smouse 1996).

322 The impact of gene flow on the persistence of natural populations is hotly debated in the
323 conservation genetics literature (Tallmon et al. 2004; Frankham et al. 2011; Bell et al. 2019). Gene
324 flow can be crucial for preventing or reversing inbreeding depression (Westemeier et al. 1998;
325 Hogg et al. 2006; Whiteley et al. 2015; Fitzpatrick et al. 2020). Gene flow can also contribute
326 critical standing variation for adaptation and may be important for future responses to climate
327 change (e.g. Taylor & Larson 2019). However, gene flow could also play a role in eliminating
328 divergence between populations and may lead to outbreeding depression (Edmands et al. 2007).
329 Our data show that coastal populations experiencing higher levels of gene flow do exhibit reduced
330 declines in genetic diversity over time. However, we also show that increased immigration into
331 the Bay Area through time likely contributed to the homogenization of outlier regions associated
332 with tidal marsh adaptation (Fig. 3, 4). Our findings contrast with a recent study on Trinidadian
333 guppies, which found that allele frequency differences in outlier loci were more resistant to gene
334 flow than non-outlier loci (Fitzpatrick et al. 2020). This study also showed that increasing gene
335 flow from main stream environments led to increased fitness of headwater guppy populations, but
336 the study was performed only for a limited time frame of ~6 generations. Our study takes a longer
337 view on the impacts of increased immigration on loci associated with local adaptation. However,
338 whether increased gene flow contributes to reduced fitness and outbreeding depression remains to
339 be seen in these tidal marsh sparrows.

340 Although we do not fully understand the contributions of the identified outlier loci to
341 fitness in tidal marshes, over 40% of these loci may contribute to increased osmoregulatory
342 performance in high salinity environments via their known roles in kidney development, function,

343 and pathology (Supplemental Table S2). Seven genes with SNPs exceeding the 99th percentile Fst
344 threshold were also previously found to be outliers in other tidal marsh sparrow species (Walsh et
345 al. 2019). These include the genes GLI family zinc finger protein 2 (GLI2) and smoothened
346 homolog (SMO), which were identified as outliers in tidal marsh Nelson's (*Ammospiza nelsoni*)
347 and Swamp Sparrows (*Melospiza georgiana*), respectively. SMO and GLI2 were both Fst outliers
348 in the *P. s. beldingi* population at Newport Bay. These genes both interact within the sonic
349 hedgehog (SHH) signaling pathway during renal development with inhibition of these genes
350 leading to disrupted kidney morphogenesis (e.g. Hu et al. 2006). Within the Bay Area, SLC9A3
351 was another Fst outlier previously found to be under selection in tidal marsh Swamp Sparrows
352 (Walsh et al. 2019). SLC9A3 plays a primary role in sodium absorption within the mammalian
353 intestine and colon (Xu et al. 2018), and has been found to be differentially expressed in different
354 fish species exposed to high salinity (Gibbons et al. 2017).

355 Among outliers identified by multiple methods, only the estrogen receptor 1 (ESR1) and
356 IgGFc-binding protein-like (FCGBP-like) genes were found to be outliers across multiple
357 populations. The ESR1 gene was previously localized to an inversion on chromosome 2 of the
358 White-throated Sparrow genome and plays a role in differential aggressive behavior between color
359 morphs in this species (Tuttle et al. 2016; Merritt et al. 2020). In addition to behavioral and
360 reproductive functions, ESR1 also contributes to the generation of important osmoregulatory
361 phenotypes matching those found in tidal marsh sparrows. First, female knockout mice lacking the
362 ESR1 locus exhibit larger kidneys (Lovegrove et al. 2004). Second, this gene is widely expressed
363 throughout the loop of Henle, collecting ducts, and nephrons in the kidney (Jelinsky et al. 2003;
364 Cheema et al. 2015). Finally, this gene inhibits Aquaporin-2 expression in the collecting duct of
365 rat and mouse kidneys with estrogen reductions leading to the production of more concentrated

366 urine (Cheema et al. 2015). This result suggests that the more concentrated urine produced by salt
367 marsh passerines (Goldstein et al. 1990; Goldstein 2006; Benham & Cheviron 2020) may be
368 related to the influence of estrogen receptor activity on AQP2 expression within the collecting
369 ducts of the kidney. Salinity and temperature are expected to increase in northern California in the
370 near future and could exacerbate osmoregulatory stress experienced by these sparrows. If the ESR1
371 and other candidate genes identified in this study contribute to important osmoregulatory
372 functions, declining allele frequencies due to gene flow may leave the Bay Area population unable
373 to adapt to the hotter and saltier tidal marshes of the future.
374
375 *Conclusions:* Research on geographic variation has long motivated the growth of natural history
376 collections. Consequently, museum collections now provide abundant material for documenting
377 the role of human activities in temporal changes to the balance between migration, drift, and
378 selection. Taking full advantage of this unique strength of collections across broad taxonomic
379 scales will be essential for developing a more robust understanding of the dynamics of plant and
380 animal population responses to ecological change. Our work demonstrates the power of analyzing
381 these replicate time series of specimens to make novel insights into how habitat loss and
382 demography interact to shape population genetic change. Specifically, the role of human-driven
383 habitat loss in altering the dynamics of gene flow and negatively impacting local adaptation will
384 be important to consider in the effective conservation of threatened populations.
385
386 **METHODS:**
387 *Sample library prep, capture, and sequencing:* Historical toepad samples were extracted using a
388 phenol-chloroform protocol in a lab space dedicated to working with degraded DNA. Historic

389 samples were treated with the USER enzyme to mitigate C->T base misincorporation (Briggs et
390 al. 2009). Modern samples were extracted using a Qiagen DNeasy extraction kit (Valencia,
391 California) and randomly sheared using a qSonica Q800R3 Sonicator (Newtown, CT). For all
392 samples, libraries were prepared for sequencing using a Kapa Illumina Hyper Prep kit
393 (supplemental methods); 500ng of DNA from each library was then pooled and hybridized in
394 solution to the SeqCap EZ probes following the manufacturer's protocols (Roche Sequencing
395 Solutions, Inc, Pleasanton, CA). Captured libraries were sequenced on two lanes of NovaSeq S4
396 and a single lane of NovaSeq SP at the University of California Berkeley Vincent J. Coates
397 Genomics Sequencing Lab.
398
399 *Population structure analyses:* Principal components analysis of the full, exon, and nongenic
400 datasets resulting from different filtering protocols (supplemental methods) was performed using
401 the python package scikit-allel v. 1.2.0 (Miles et al. 2020). For all SNP datasets, we removed
402 singletons and pruned SNPs that were in linkage (r=0.25) in 2500 bp blocks. The estimated
403 effective population size (Ne) in DyStruct was set to 100,000 after preliminary analyses showed
404 that results were robust to varying the Ne setting from the smaller (25,000) to larger population
405 sizes (500,000) previously reported for this species (Benham & Cheviron 2019). Generation time
406 was set to 2.5 years (Bird et al. 2020) and we binned samples into eight 10-year windows spanning
407 48 generations from 1890 (generation 0) to 2017 (generation 48). We ran DyStruct three times
408 independently for K = 1 to 12, where K is the number of pre-defined population clusters. The most
409 likely value of K was identified using a cross-validation, hold-out method within the program.
410

411 *Genetic diversity and demographic analyses:* We compared historical and modern patterns of
412 migration and genetic diversity across California with Estimated Effective Migration Surfaces in
413 the program EEMS (Petkova et al. 2016). For both the modern and historic datasets, input files
414 were generated using the bed2diffs function in EEMS. We ran three independent runs of the
415 MCMC chain. Each chain started from a different random seed and ran for 2 million iterations of
416 burn-in followed by 5 million iterations. A population grid density of 250 demes was selected and
417 we adjusted the qEffctProposalS2 (0.01) and qSeedsProposalS2 (0.2) settings upwards to ensure a
418 proposal acceptance rate between 10-40%. We used the R package rEEMSplots to assess model
419 fit and convergence among the three independent chains.

420 To match sequence depth between modern and historical samples, population genetic
421 statistics were estimated using a dataset with downsampled coverage in modern samples and the
422 removal of C->T, G->A mutations. We estimated Tajima's D, nucleotide diversity (π), and
423 Watterson's Theta (θ) for each exon (n=9,971 loci) and non-genic target region (n=1,024 loci) in
424 the dataset using scikit-allel v. 1.2.0 (Miles et al. 2020). Heterozygosity and the inbreeding
425 coefficient for each individual were estimated for nongenic and exon datasets in VCFtools
426 (Danecek et al. 2011) using the --het function. These metrics were estimated for each population
427 (defined by population genetic clusters identified using PCA analysis, n=5) and sampling time
428 (modern vs. historic). We also estimated these metrics across three and six time points in Newport
429 Bay (*P. s. beldingi*) and San Francisco Bay (*P. s. alaudinus*), respectively. To estimate the
430 magnitude of temporal change in each metric we subtracted the historic from the modern value.
431 Within each population we assessed whether significant differences existed between sampling
432 points using ANOVA and post-hoc TukeyHSD tests or t-tests depending on the number of time

433 points sampled. All statistical analyses were performed in the open source program R version 3.5.1
434 (https://www.r-project.org/).
435 We used the program *Moments* (Jouganous et al. 2017) to assess temporal change in
436 patterns of gene flow among California populations of the Savannah sparrow. We fit an isolation
437 with migration model to the three dimensional site frequency spectrum of eastern California (*P. s.*
438 *nevadensis*), Bay Area (*P. s. alaudinus*), and Newport Bay (*P. s. beldingi*) birds. We fit this model
439 to the site frequency spectrum of a historic (26,080 SNPs) and modern (21,758 SNPs) dataset of
440 individuals from these populations. See supplemental methods for further details on data filtering
441 and program parameterization.
442
443 *Spatial and temporal selection analyses:* We identified putative outlier loci using the
444 downsampled dataset with C->T and G->A sites removed, a maf filter of 0.05, and only retained
445 sites with no missing data. We divided the dataset into five populations: (1) freshwater-adapted
446 populations of eastern California (*P. s. nevadensis*); (2) Humboldt Bay estuary (*P. s. alaudinus*);
447 (3) San Francisco Bay estuary (*P. s. alaudinus*); (4) Morro Bay estuary (*P. s. alaudinus/beldingi*);
448 and (5) Newport Bay estuary (*P. s. beldingi*). Each of these five populations were further divided
449 into historic (sampled 1900-1920) and modern (sampled after 2000) datasets. First, we estimated
450 Weir & Cockerham's Fst divergence between each estuary population (pops 2-4) and inland
451 localities across eastern California (pop 1) for both the historical and modern datasets. SNPs were
452 identified as being potentially under selection within tidal marsh environments if they exceeded
453 the 99th percentile of Fst. Second, we estimated Hudson's Fst (Bhatia et al. 2012) and Dxy between
454 salt and freshwater-adapted populations for 5Kb windows with a step size of 2.5 Kb following
455 approaches used to analyze other exome datasets (e.g., Schweizer et al. 2021). Third, we employed

456 a Latent Factor Mixed Models approach (LFMM; Frichot et al. 2013) to identify SNPs with
457 significant associations with mean salinity estimates from a previously published dataset (Benham
458 & Bowie 2021). Loci were designated as outliers if the 5kb region was in the top 5% of Fst
459 divergence, top 5% of Dxy divergence, and contained a SNP with a p-value <0.05 based on LFMM
460 analyses for each of the four populations. For both per SNP Fst estimates and 5kb windows we
461 compared temporal changes in Fst divergence by subtracting historical Fst from modern Fst
462 divergence in each locus to estimate ΔFst. We used t-tests to compare whether outlier versus non-
463 outlier regions exhibited significantly different changes in Fst through time.

464 To assess the degree of correlation between historical deviations in mean allele frequency
465 and the direction of allele frequency change through time. We estimated allele frequencies using
466 VCFtools. We classified SNPs as outliers if they exceeded the 99th percentile of Fst divergence.
467 We compared Pearson's correlation of the outlier SNPs to non-outlier SNPs by performing a
468 permutation test where we randomly sampled an equal number of non-outlier to outlier SNPs and
469 calculated correlations 1000 times. We then compared whether the non-outlier loci were both
470 statistically different from 0 (no correlation) or the correlation estimated for the outlier loci. We
471 used changes in the proportion of eastern California ancestry found in each of the four coastal
472 populations from the DyStruct analysis to approximate temporal increases in gene flow.

473

474 **ACKNOWLEDGEMENTS:**

475 We thank the following museums and individuals for providing access to their collections for
476 this study: Museum of Vertebrate Zoology (Carla Cicero); California Academy of Sciences
477 (Maureen Flannery, Jack Dumbacher); San Diego Museum of Natural History (Philip Unitt); Los
478 Angeles County Museum of Natural History (Allison Shultz, Kimball Garrett); American

Museum of Natural History (Paul Sweet, Tom Trombone); Field Museum of Natural History (John Bates, Shannon Hackett, Ben Marks); University of Washington Burke Museum (Sharon Burkes); University of Montana Zoological Museum (Angela Hornesby). Personnel at the California Department of Fish and Wildlife, California State Parks, and United States Fish and Wildlife Service assisted with access to federal and state lands. Finally we thank Elizabeth Beckman, Ammon Corl, Mackenzie Kirchner-Smith, Jackie Childers, Rosa Jimenez, Dan Wait, and Cynthia Wang-Claypool for useful comments on the manuscript. This study was supported by an NSF PRFB fellowship (#1812282) to PMB.

REFERENCES:

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Bell, D. A., Robinson, Z. L., Funk, W. C., Fitzpatrick, S. W., Allendorf, F. W., Tallmon, D. A., & Whiteley, A. R. (2019). The Exciting Potential and Remaining Uncertainties of Genetic Rescue. *Trends in Ecology and Evolution*, *34*(12), 1070–1079. https://doi.org/10.1016/j.tree.2019.06.006

Benham, P. M., & Bowie, R. C. K. (2021). The influence of spatially heterogeneous anthropogenic change on bill size evolution in a coastal songbird. *Evolutionary Applications*, (September), 1–18. https://doi.org/10.1111/eva.13144

Benham, P. M., & Cheviron, Z. A. (2019). Divergent mitochondrial lineages arose within a large, panmictic population of the Savannah sparrow (*Passerculus sandwichensis*) . *Molecular Ecology*, 1765–1783. https://doi.org/10.1111/mec.15049

Benham, P. M., & Cheviron, Z. A. (2020). Population history and the selective landscape shape patterns of osmoregulatory trait divergence in tidal marsh Savannah sparrows (Passerculus sandwichensis). *Evolution*, *74*(1). https://doi.org/10.1111/evo.13886

Bhatia, G., Patterson, N., Sankararaman, S., & Price, A. L. (2013). Estimating and interpreting FST: The impact of rare variants. *Genome Research*, *23*(9), 1514–1521. https://doi.org/10.1101/gr.154831.113

Bi, K., Linderoth, T., Vanderpool, D., Good, J. M., Nielsen, R., & Moritz, C. (2013). Unlocking the vault: Next-generation museum population genomics. *Molecular Ecology*, *22*(24), 6018–6032. https://doi.org/10.1111/mec.12516

Bi, K., Linderoth, T., Singhal, S., Vanderpool, D., Patton, J. L., Nielsen, R., … Good, J. M. (2019). Temporal genomic contrasts reveal rapid evolutionary responses in an alpine mammal during recent climate change. *PLOS Genetics*, *15*(5), e1008119. https://doi.org/10.1371/journal.pgen.1008119

Billerman, S. M., & Walsh, J. (2019). Historical DNA as a tool to address key questions in avian biology and evolution: A review of methods, challenges, applications, and future directions. *Molecular Ecology Resources*, (July), 1755-0998.13066. https://doi.org/10.1111/1755-0998.13066

Bird, J. P., Martin, R., Akçakaya, H. R., Gilroy, J., Burfield, I. J., Garnett, S. T., … Butchart, S. H. M. (2020). Generation lengths of the world's birds and their implications for extinction risk. *Conservation Biology*, *00*(0), 1–10. https://doi.org/10.1111/cobi.13486

Briggs, A. W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., & Pääbo, S. (2009). Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Research*, *38*(6), 1–12. https://doi.org/10.1093/nar/gkp1163

Cade, T., & Bartholomew, G. A. (1959). Sea-water and salt utilization by Savannah sparrows. *Physiological Zoology*, *32*(4), 230–238. Retrieved from http://www.jstor.org/stable/30155400 .

Cheema, M. U., Irsik, D. L., Wang, Y., Miller-Little, W., Hyndman, K. A., Marks, E. S., … Norregaard, R. (2015). Estradiol regulates AQP2 expression in the collecting duct: A novel inhibitory role for estrogen receptor α. *American Journal of Physiology - Renal Physiology*, *309*(4), F305–F317. https://doi.org/10.1152/ajprenal.00685.2014

Currat, M., Ruedi, M., Petit, R. J., & Excoffier, L. (2008). The hidden side of invasions: Massive introgression by local genes. *Evolution*, *62*(8), 1908–1920. https://doi.org/10.1111/j.1558-5646.2008.00413.x

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dias, P. C. (1996). Sources and sinks in population biology. *Trends in Ecology and Evolution*, *11*(8), 326–330. https://doi.org/10.1016/0169-5347(96)10037-9

Edmands, S. (2007). Between a rock and a hard place: Evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology*, *16*(3), 463–475. https://doi.org/10.1111/j.1365-294X.2006.03148.x

Feng, S., Fang, Q., Barnett, R., Li, C., Han, S., Kuhlwilm, M., … Zhang, G. (2019). The Genomic Footprints of the Fall and Recovery of the Crested Ibis. *Current Biology*, *29*(2), 340-349.e7. https://doi.org/10.1016/j.cub.2018.12.008

Fitton, S.D. (2008). Bryant's Savannah Sparrow (*Passerculus sandwichensis alaudinus*). Shuford, W.D. & Gardali, T., eds. California Bird Species of Special Concern: A ranked assessment of species, subspecies, and distinct populations of birds of immediate conservation concern in California. *Studies of Western Birds 1.* Western Field Ornithologists, Camarillo, California, and California Department of Fish and Game, Sacramento.

Fitzpatrick, S. W., Bradburd, G. S., Kremer, C. T., Salerno, P. E., Angeloni, L. M., & Funk, W. C. (2020). Genomic and Fitness Consequences of Genetic Rescue in Wild Populations. *Current Biology*, *30*(3), 517-522.e5. https://doi.org/10.1016/j.cub.2019.11.062

Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, *126*(2), 131–140. https://doi.org/10.1016/j.biocon.2005.05.002

Frankham, R., Ballou, J. D., Eldridge, M. D. B., Lacy, R. C., Ralls, K., Dudash, M. R., & Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology*, *25*(3), 465–475. https://doi.org/10.1111/j.1523-1739.2011.01662.x

Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, *30*(7), 1687–1699. https://doi.org/10.1093/molbev/mst063

Gaggiotti, O. E., & Smouse, P. E. (1996). Stochastic Migration and Maintenance of Genetic Variation in Sink Populations. *American Naturalist*, *147*(6), 919–945.

Gerdes, G.L., Primbs, E.R.J., Browning, B.M. 1974. Natural resources of Morro Bay. Their status and future. State of California, Department of Fish and Game, Coastal Wetland Series #8.

Gibbons, T. C., Metzger, D. C. H., Healy, T. M., & Schulte, P. M. (2017). Gene expression plasticity in response to salinity acclimation in threespine stickleback ecotypes from different salinity habitats. *Molecular Ecology*, *26*(10), 2711–2725. https://doi.org/10.1111/mec.14065

Goldstein, D. L. (2006). Osmoregulatory biology of saltmarsh passerines. *Terrestrial Vertebrates of Tidal Marshes: Evolution, Ecology, and Conservation*, (32), 110–118.

Goldstein, D. L., Williams, J. B., & Braun, E. J. (1990). Osmoregulation in the field by salt-marsh savannah sparrows Passerculus sandwichensis beldingi. *Physiological Zoology*, *63*(4), 669–682.

576 Gompert, Z., Springer, A., Brady, M., Chaturvedi, S., & Lucas, L. K. (2021). Genomic time-
577 series data show that gene flow maintains high genetic diversity despite substantial genetic
578 drift in a butterfly species. *Molecular Ecology*, (April), 1–18.
579 https://doi.org/10.1111/mec.16111

580 Grinnell, J., Miller, A.H. (1944). The distribution of the birds of California. Pacific Coast
581 Avifauna 27: 1-608.

582 Hogg, J. T., Forbes, S. H., Steele, B. M., & Luikart, G. (2006). Genetic rescue of an insular
583 population of large mammals. *Proceedings of the Royal Society B: Biological Sciences*,
584 *273*(1593), 1491–1499. https://doi.org/10.1098/rspb.2006.3477

585 Holmes, M. W., Hammond, T. T., Wogan, G. O. U., Walsh, R. E., Labarbera, K., Wommack, E.
586 A., … Nachman, M. W. (2016). Natural history collections as windows on evolutionary
587 processes. *Molecular Ecology*, *25*(4), 864–881. https://doi.org/10.1111/mec.13529

588 Hu, M. C., Mo, R., Bhella, S., Wilson, C. W., Chuang, P. T., Hui, C. C., & Rosenblum, N. D.
589 (2006). GLI3-dependent transcriptional repression of Gli1, Gli2 and kidney patterning
590 genes disrupts renal morphogenesis. *Development*, *133*(3), 569–578.
591 https://doi.org/10.1242/dev.02220

592 Jelinsky, S. A., Harris, H. A., Brown, E. L., Flanagan, K., Zhang, X., Tunkey, C., … Evans, M.
593 J. (2003). Global transcription profiling of estrogen activity: Estrogen receptor α regulates
594 gene expression in the kidney. *Endocrinology*, *144*(2), 701–710.
595 https://doi.org/10.1210/en.2002-220728

Joseph, T. A., & Pe'er, I. (2019). Inference of Population Structure from Time-Series Genotype Data. *American Journal of Human Genetics*, *105*(2), 317–333. https://doi.org/10.1016/j.ajhg.2019.06.002

Jouganous, J., Long, W., Ragsdale, A. P., & Gravel, S. (2017). Inferring the Joint Demographic History of Multiple Populations: Beyond the Diffusion Approximation. *Genetics*, *206*(3), 1549–1567. https://doi.org/10.1534/genetics.117.200493

LaDochy, S., Medina, R., & Patzert, W. (2007). Recent Calfornia climate variability: spatial and temporal patterns in temperature trends. *Climate Research*, *33*, 159–169.

Lebassi, B., Gonzalez, J., Fabris, D., Maurer, E., Miller, N., Milesi, C., … Bornstein, R. (2009). Observed 1970-2005 cooling of summer daytime temperatures in coastal California. *Journal of Climate*, *22*(13), 3558–3573. https://doi.org/10.1175/2008JCLI2111.1

Lino, A., Fonseca, C., Rojas, D., Fischer, E., & Ramos Pereira, M. J. (2019). A meta-analysis of the effects of habitat loss and fragmentation on genetic diversity in mammals. *Mammalian Biology*, *94*, 69–76. https://doi.org/10.1016/j.mambio.2018.09.006

Lovegrove, A. S., Sun, J., Gould, K. A., Lubahn, D. B., Korach, K. S., Lane, P. H., … Estrogen, P. H. L. (2004). How Genes Influence Renal Function and Disease Estrogen receptor ␣ -mediated events promote sex-specific diabetic glomerular hypertrophy. *Blood*, *2196*, 586–591.

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M. P., Elston, D. A., Scott, E. M., … Watt, A. D. (2010). Long-term datasets in biodiversity research and monitoring: Assessing

change in ecological communities through time. *Trends in Ecology and Evolution*, *25*(10), 574–582. https://doi.org/10.1016/j.tree.2010.06.016

Marshall, J., & Dedrick, K. (1994). Endemic Song Sparrows and Yellowthroats of San Francisco Bay. *Studies in Avian Biology*, *15*(15), 316-327.

Merritt, J. R., Grogan, K. E., Zinzow-Kramer, W. M., Sun, D., Ortlund, E. A., Yi, S. V., & Maney, D. L. (2020). A supergene-linked estrogen receptor drives alternative phenotypes in a polymorphic songbird. *Proceedings of the National Academy of Sciences*, *117*(35), 202011347. https://doi.org/10.1073/pnas.2011347117

Miles, A. et al. 2020 Scikit-Allel: https://zenodo.org/record/4759368#.YQ2eqC2cYWo

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*(1), 94–100. https://doi.org/10.1038/ng.3464

Poulson, T. L., & Bartholomew, G. A. (1962). Salt balance in the Savannah Sparrow. *Physiological Zoology*, *35*, 109–119.

Powell, A. N. 2006. Are southern California's fragmented saltmarshes capable of sustaining endemic bird populations? *Studies in Avian Biology*, (32), 198-204.

Sánchez-Barreiro, F., Gopalakrishnan, S., Ramos-Madrigal, J., Westbury, M. V., de Manuel, M., Margaryan, A., … Gilbert, M. T. P. (2021). Historical population declines prompted significant genomic erosion in the northern and southern white rhinoceros (Ceratotherium simum). *Molecular Ecology*, (June), 1–15. https://doi.org/10.1111/mec.16043

Schweizer, R. M., Jones, M. R., Bradburd, G. S., Storz, J. F., Senner, N. R., Wolf, C., & Cheviron, Z. A. (2021). Broad Concordance in the Spatial Distribution of Adaptive and Neutral Genetic Variation across an Elevational Gradient in Deer Mice. *Molecular Biology and Evolution*, *38*(10), 4286–4300. https://doi.org/10.1093/molbev/msab161

Suchan, T., Pitteloud, C., Gerasimova, N. S., Kostikova, A., Schmid, S., Arrigo, N., … Alvarez, N. (2016). Hybridization capture using RAD probes (hyRAD), a new tool for performing genomic analyses on collection specimens. *PLoS ONE*, *11*(3), 1–22. https://doi.org/10.1371/journal.pone.0151651

Takekawa, J. Y., Woo, I., Spautz, H., Nur, N., Letitia Grenier, J., Malamud-Roam, K., … Wainwright-De La Cruz, S. E. (2006). Environmental threats to tidal-marsh vertebrates of the San Francisco Bay estuary. *Studies in Avian Biology*, (32), 176–197.

Tallmon, D. A., Luikart, G., & Waples, R. S. (2004). The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology and Evolution*, *19*(9), 489–496. https://doi.org/10.1016/j.tree.2004.07.003

Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology & Evolution*, *3*(2), 170–177. https://doi.org/10.1038/s41559-018-0777-y

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., … Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications*, *9*(7), 892–908. https://doi.org/10.1111/eva.12367

656 Tuttle, E. M., Bergland, A. O., Korody, M. L., Brewer, M. S., Newhouse, D. J., Minx, P., …
657 Balakrishnan, C. N. (2016). Divergence and functional degradation of a sex chromosome-
658 like supergene. *Current Biology*, *26*(3), 344–350. https://doi.org/10.1016/j.cub.2015.11.069

659 U.S. Census Bureau 2019, Population Estimates Program (PEP), updated annually.
660 https://www.census.gov/programs-surveys/popest.html

661 van der Valk, T., Díez-del-Molino, D., Marques-Bonet, T., Guschanski, K., & Dalén, L. (2019).
662 Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in
663 Eastern Gorillas. *Current Biology*, *29*(1), 165-170.e6.
664 https://doi.org/10.1016/j.cub.2018.11.055

665 Walsh, J., Benham, P. M., Deane-Coe, P. E., Arcese, P., Butcher, B. G., Chan, Y. L., … Lovette,
666 I. J. (2019). Genomics of rapid ecological divergence and parallel adaptation in four tidal
667 marsh sparrows. *Evolution Letters*, 1–15. https://doi.org/10.1002/evl3.126

668 Wandeler, P., Hoeck, P. E. A., & Keller, L. F. (2007). Back to the future: museum specimens in
669 population genetics. *Trends in Ecology and Evolution*.
670 https://doi.org/10.1016/j.tree.2007.08.017

671 Welch, A. J., Wiley, A. E., James, H. F., Ostrom, P. H., Stafford, T. W., & Fleischer, R. C.
672 (2012). Ancient DNA reveals genetic stability despite demographic decline: 3,000 years of
673 population history in the endemic hawaiian petrel. In *Molecular Biology and Evolution*
674 (Vol. 29, pp. 3729–3740). https://doi.org/10.1093/molbev/mss185

Westemeier, R. L., Brawn, J. D., Simpson, S. A., Esker, T. L., Jansen, R. W., Walk, J. W., … Paige, K. N. (1998). Tracking the long-term decline and recovery of an isolated population. *Science*, *282*(5394), 1695–1698. https://doi.org/10.1126/science.282.5394.1695

Wheelwright, N.T., Rising, J.D. (2008). Savannah Sparrow (*Passerculus sandwichensis*). *in* Birds of North American Online (A. Poole, Ed.). Cornell Lab of Ornithology, Ithaca, NY.

Whiteley, A. R., Fitzpatrick, S. W., Funk, W. C., & Tallmon, D. A. (2015). Genetic rescue to the rescue. *Trends in Ecology and Evolution*, *30*(1), 42–49. https://doi.org/10.1016/j.tree.2014.10.009

Xu, H., Ghishan, F. K., & Kiela, P. R. (2018). SLC9 gene family: Function, expression, and regulation. *Comprehensive Physiology*, *8*(2), 555–583. https://doi.org/10.1002/cphy.c170027

Zembal, R. & Hoffmann, S.M. (2010). A survey of the Belding's Savannah sparrow (*Passerculus sandwichensis beldingi*) in California 2010. Calif. Dep. Fish and Game, Wildlife Branch, Nongame Wildlife Program Report 2010-10, Sacramento, CA 17 pp.

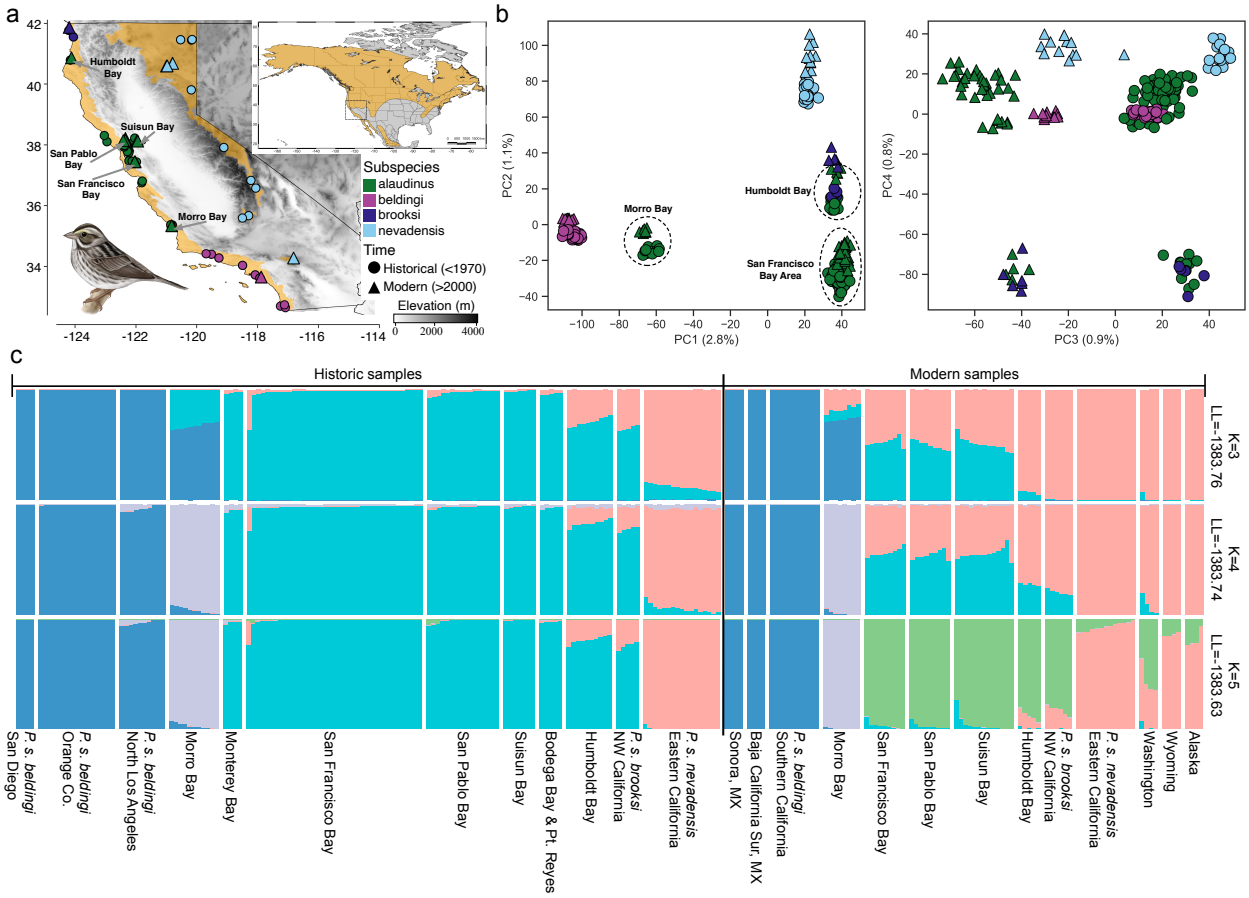

Figure 1: Spatial and temporal patterns of population structure within California populations of the Savannah sparrow (*Passerculus sandwichensis*). **(a)** California breeding distribution of the Savannah sparrow (ochre). Circles represent historical (pre-1970) sampling localities of specimens sequenced in the study, and triangles represent modern (post-2000) sampling localities. Inset shows extent of the entire North American breeding distribution of the Savannah sparrow in ochre. **(b)** Principal components analysis of the full exome capture dataset (60,765 SNPs). Left panel shows PC1 (3.1% variance explained) versus PC2 (1.1% variance explained) and right panel shows PC3 (0.9% variance explained) versus PC4 (0.8% variance explained). Colors correspond to different subspecies as illustrated in Fig. 1a. Dashed circles denote different clusters corresponding to different populations within *P. s. alaudinus.* **(c)** Results from DyStruct analysis. Runs with clusters specified as K=3-5 all exhibited very similar log-likelihoods and are illustrated here. Colors correspond to different K clusters. Artwork of *P. s. beldingi* individual courtesy of Jillian Nichol Ditner

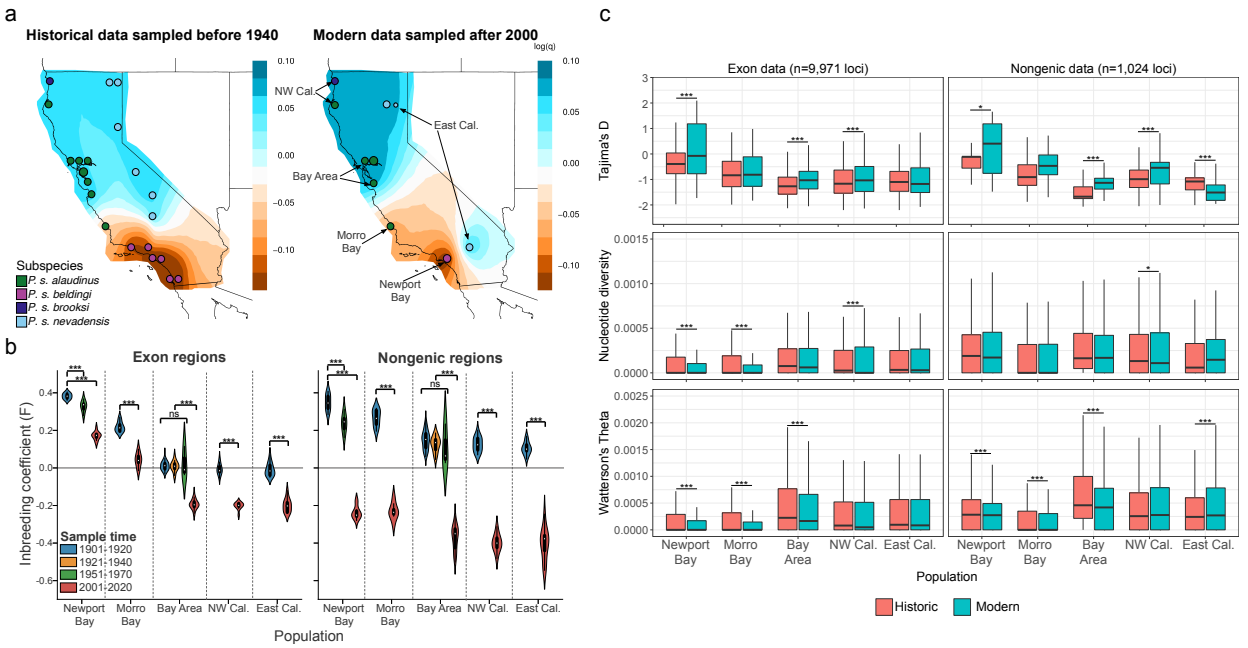

Figure 2: Spatial and temporal patterns of genetic diversity within California Savannah sparrows. **(a)** Historical and modern variation in genetic diversity across California estimated using the Estimated Effective Migration Surfaces (EEMS) program. Contour plots vary from low (dark brown) to high (dark blue) posterior mean diversity rates. **(b)** Variation in the inbreeding coefficient across different populations through time. The five populations correspond to the five genetic clusters shown in Fig. 1b. Data were divided into exon regions and nongenic regions >100 kb from any exon in the genome. Colors correspond to bins of different sampling times. **(c)** Boxplots comparing historical (red) and modern (blue) estimates of Tajima's D, Watterson's theta, and nucleotide diversity across the five different populations for both exon (left panel) and nongenic (right panel) target regions. Outlier regions removed from plot to simplify visualization. For both **b** & **c** plots are based on a downsampled dataset with C->T and G->A mutations removed. Asterisks in **b** & **c** denote p-values based on t-tests or ANOVA among sampling time groups (ns: non-significant; *p<0.05; **p<0.01; ***: p<0.0001). In **c** only historical modern comparisons with significant differences are denoted.

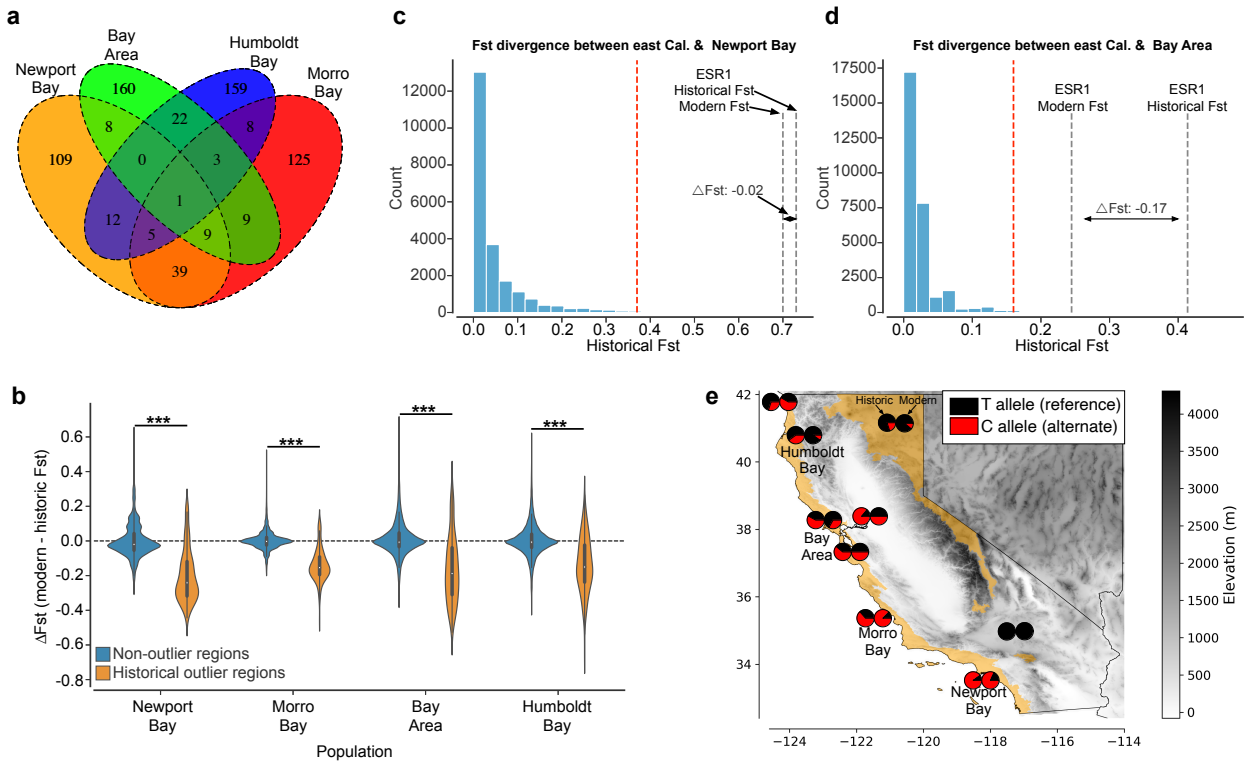

Figure 3: Temporal variation in Fst divergence. **(a)** Venn diagram showing overlap in genes with outlier SNPS across the four tidal marsh populations. **(b)** Change in Fst divergence between eastern California and four tidal marsh populations of Savannah sparrow. ΔFst for each SNP was estimated as the difference between historic and modern Fst divergence with positive values indicating increasing Fst divergence through time. SNPs above the 99th percentile of divergence were identified as historical outlier SNPs. Asterisks denote p-values from t-tests of ΔFst between outlier and non-outlier SNPs (ns: non-signifiant; *p<0.05; **p<0.01; ***: p<0.0001). **(c-d)** Distribution of per SNP Fst estimates for historical comparison between Newport Bay (*P. s. beldingi*) and Bay Area (*P. s. alaudinus*) with eastern California (*P. s. nevadensis*). Red dashed line denotes the mean 99th percentile threshold for calling outlier SNPs. Gray dashed lines represent modern and historic Fst estimates for a C->T mutation in the 3' UTR region of the final exon of the Estrogen Receptor 1 gene (ESR1). **(e)** Spatial variation in allele frequencies for the ESR1 C->T SNP across nine sampling localities with paired modern and historic sampling. Pie charts on left illustrate historical allele frequencies and pie charts on the right modern allele frequencies.

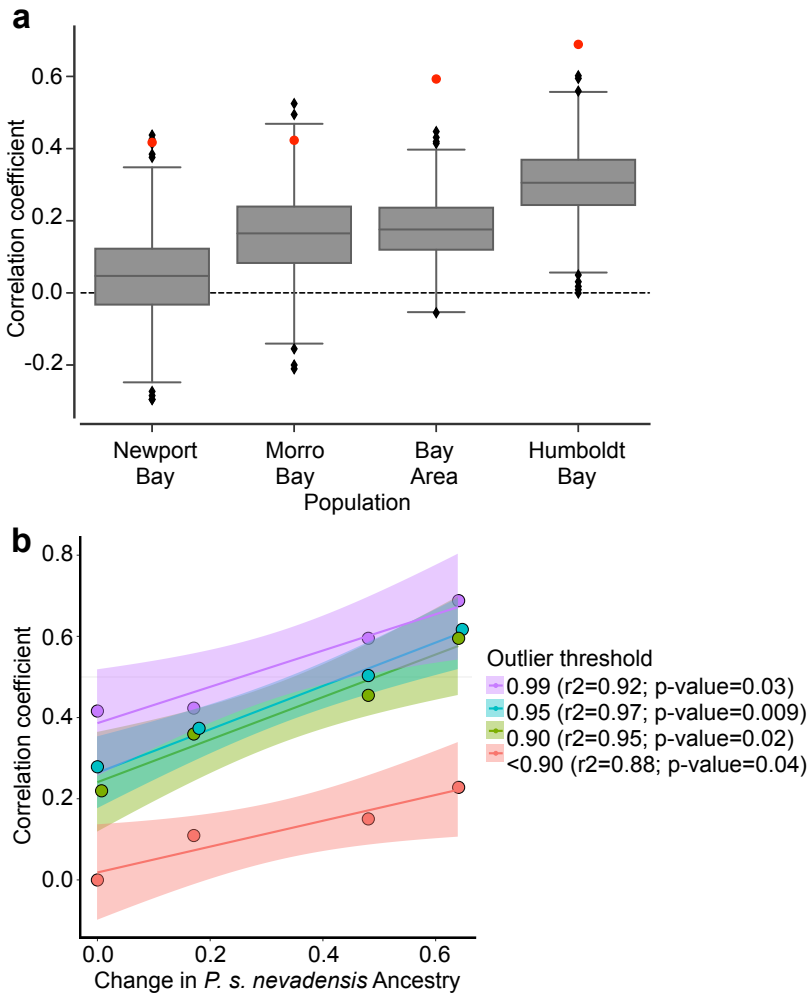

Figure 4: (a) Correlations between the historical deviation from mean allele frequency and direction of temporal change in each tidal marsh population. More positive correlations indicate greater changes in allele frequency towards the mean due to higher levels of gene flow. Red circles show observed correlation for outlier SNPs and gray box plots indicate the null distribution from 1000 permutations of non-outlier SNPs. **(c)** Regression analyses showing relationship between change in the proportion of ancestry from *P. s. nevadensis* populations in each tidal marsh population and the correlation coefficient for outlier SNPs shown in (b). Ancestry proportions from DyStruct analysis (Fig. 1c). Different colors represent SNPs inferred to be outliers from different percentile thresholds with red points showing data from non-outlier SNPs.

Table 1: Mean Fst divergence between eastern California and each of four tidal marsh populations of the Savannah sparrow in both the historic and modern datasets.

Population	historic sample size	modern sample size	Number SNPs	Historic mean Fst (sd)	Modern mean Fst (sd)
Newport Bay	10	11	4495	0.098 (0.008)	0.099 (0.009)
Morro Bay	11	8	5408	0.054 (0.003)	0.062 (0.006)
Bay Area	31	37	12021	0.015 (0.001)	0.013 (0.001)
Humboldt Bay	10	5	6466	0.014 (0.001)	-0.002 (0.002)