Published in final edited form as:
Curr Biol. 2025 March 24; 35(6): 1391–1399.e6. doi:10.1016/j.cub.2025.01.058.

Translocations spur population growth but fail to prevent genetic erosion in imperiled Florida Scrub-Jays

Tyler Linderoth[1,8,9,*], **Lauren Deaner**[2], **Nancy Chen**[3], **Reed Bowman**[4], **Raoul K. Boughton**[5], **Sarah W. Fitzpatrick**[1,6,7]

[1]Michigan State University, W.K. Kellogg Biological Station, 3700 East Gull Lake Drive, Hickory Corners, MI 49060, USA

[2]Verdantas, 8306 Laurel Fair Circle, Suite 120, Tampa, FL 33610, USA

[3]University of Rochester, Department of Biology, 477 Hutchison Hall, P.O. Box 270211, Rochester, NY 14627, USA

[4]Archbold Biological Station, Avian Ecology Program, 123 Main Drive, Venus, FL 33960, USA

[5]Mosaic Fertilizer LLC, Land and Resource Management, 13830 Circa Crossing Drive, Lithia, FL 33547, USA

[6]Michigan State University, Department of Integrative Biology, 288 Farm Lane, East Lansing, MI 48824, USA

[7]Michigan State University, Ecology, Evolution, and Behavior Program, East Lansing, MI 48824, USA

[8]X (formerly Twitter): @tpLinderoth

[9]Lead contact

SUMMARY

Land and natural resource use in addition to climate change can restrict populations to degraded and fragmented habitats, catalyzing extinction through the reinforced interplay of small population size and genetic decay. Translocating individuals is a powerful strategy for overcoming direct threats from human development and reconnecting isolated populations but is not without risks.[1] Habitat Management Plan analyses under section 7 of the U.S. Endangered Species Act determined that multiple subpopulations of Federally Threatened Florida Scrub-Jays (*Aphelocoma coerulescens*, hereafter FSJ) belonging to a metapopulation on Florida's west coast were declining demographic sinks, occupying areas where agriculture and fire suppression had degraded and

*Correspondence: lindero1@msu.edu.

AUTHOR CONTRIBUTIONS

T.L., S.W.F., R.K.B., R.B., L.D., and N.C. designed and performed research and wrote the paper. L.D., R.K.B., and R.B. undertook field collections and data management. L.D., R.K.B., R.B., and T.L. performed data quality control. T.L. and L.D. analyzed data. T.L. contributed software and new analytical tools.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

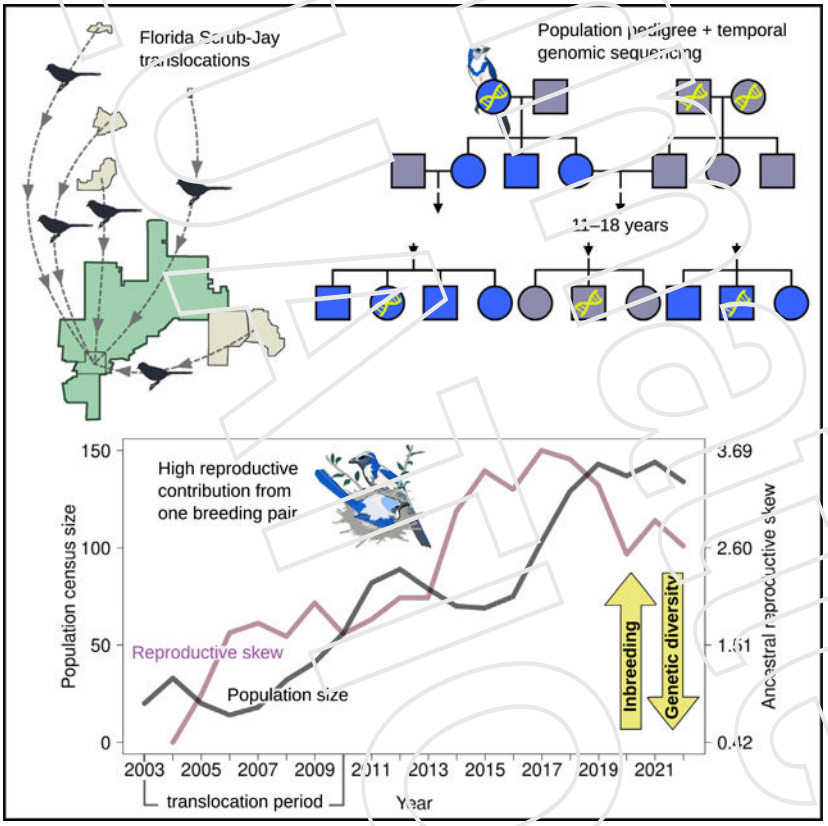
Supplemental information can be found online at https://doi.org/10.1016/j.cub.2025.01.058.

fragmented the habitat.[2] In order to increase the viability of the overall metapopulation, 51 FSJs from five of these small subpopulations in areas to be mined were translocated throughout 2003–2019 into a larger site of more contiguous, recently restored habitat at the core of the metapopulation, which contained a small resident population.[3] Prior to translocations and for nearly two decades afterward, this core population, referred to as the M4 core region (CR) population, was extensively monitored, yielding a nearly complete pedigree. We used this pedigree, along with temporal genomic analyses and simulations, to show that translocations coupled with habitat restoration generated rapid population growth, but high reproductive skew increased inbreeding and led to genetic erosion. This mechanistic understanding of mixed conservation outcomes highlights the importance of monitoring and the potential need for genetic rescue to offset consequences of reproductive skew following translocations, regardless of demographic recovery.

In brief

Linderoth et al. find that translocations of Florida Scrub-Jays aid the recovery of the species by catalyzing population growth but that reproductive skew dampens accompanying genetic benefits, ultimately leading to increased inbreeding and genetic decay. These results demonstrate the potential need for genetic rescue even in growing populations.

Graphical abstract



RESULTS

Translocations fuel demographic recovery

Thirteen individuals comprising four family groups inhabited the metapopulation 4 (M4) core region (CR) (Figures 1A and 1B) prior to translocations. These individuals and their immediate offspring were considered CR "residents" ($n = 23$). During 2003–2005 and 2007–2010, 51 Florida Scrub-Jays (FSJs) from five at-risk subpopulations in areas to be mined (Figure 1A) were translocated to an area of restored habitat within the CR (Figures 1B and 1C). The first four translocations had limited settlement success, so methodologies were adapted to time translocations with natural FSJ dispersal patterns, resulting in improved settlement for translocations conducted in 2008–2010. Following each of the modified translocations, the CR population began to increase in size (Figure 1C), while also expanding geographically as new family groups established territories throughout restored habitat in the CR (Figure 1B). The CR population continued to grow rapidly years after the last jays were translocated, such that by 2022 the population had increased 10-fold, showing a clear positive demographic response to translocations. We refer to the CR population in years 2021 and on as the "contemporary" population.

Reproductive skew revealed by the population pedigree

High reproductive variance among 74 founding individuals of the contemporary CR population (residents plus translocated jays) (Table S1) and their descendants led to a decrease in the number of founding lineages present in the population over time (Figures 2A and 2B), with a bias toward contributions from translocated lineages to annual cohorts starting approximately halfway through the translocation period (Figure 1C). Based on the nearly complete population pedigree, 16 individuals (hereafter collectively called the "top 16" ancestors) contributed over 76% of the population's founding ancestry, i.e., the genetic material originating from resident and translocated individuals, since 2015, which reached 88% by 2022 (Figure 2).

The pedigree also revealed that one breeding pair consisting of a translocated male from site 13, RLS-K, and a translocated female from site 1, WSA-K, made exceptionally high genetic contributions to the CR population (Figure 2; Table S1). This pair had consistently high nesting success, fledging all offspring in five of six nest attempts. On average, since their pairing in 2008, they collectively accounted for 24% (SD = 0.06) of the expected CR population founder ancestry, with a maximum contribution of 32% in 2017, 4 years after their presumed deaths, indicating successful reproduction of their offspring. From 2008 to 2022, the average genetic contributions by RLS-K and WSA-K to the CR population were, respectively, at least 2.43 and 1.83 higher than that of any other resident or translocated individual.

Erosion of genome-wide variation from reproductive skew

Higher inbreeding is an inevitable outcome of increased reproductive skew that can be particularly pronounced in small, isolated populations. We assessed the genetic consequences of reproductive skew in the CR population by sequencing the genome of one nestling from each 2021 CR breeding pair (contemporary sample, $n = 28$), which we

compared with the genomes of 29 translocated jays and 16 residents. We also sequenced the genomes of 13 jays from five other M4 subpopulations (Figure 1A) and identified 2,049,176 high-quality SNPs segregating across all individuals. The average sequencing depth per individual was 7.23 (SD = 1.0). The donor populations were collectively genetically separated from the resident CR population by an F_{ST} of 3%. The contemporary population diverged by 2.6% from the resident CR population and 2.4% from the translocated jays in only around 3–5 generations. In the space defined by the first two principal components (PCs) of a principal-component analysis (PCA) of genetic variation across geographic space and time, contemporary individuals gravitated toward translocated jays with high reproductive contributions (Figure 3A).

In line with a partial genomic sweep from reproductive skew, individual inbreeding estimated from genomic data, F, was significantly higher among contemporary individuals ($\overline{F}_{contemporary} = 0.015$, SD = 0.025) compared with residents ($\overline{F}_{resident} = 0.010$, SD = 0.023, Mann-Whitney U = 313, two-tailed test, $p = 1.046 \times 10^{-4}$; Figure 3B), while average heterozygosity was significantly lower among contemporary individuals ($\overline{H}_{contemporary} = 6.41 \times 10^{-4}$, SD = 6.01×10^{-5}) compared with residents ($\overline{H}_{resident} = 9.22 \times 10^{-4}$, SD = 1.74×10^{-4}, $t(13.35) = -5.662$, two-tailed test, $p = 7.018 \times 10^{-5}$; Figure 3C). Inbreeding was also significantly higher and average heterozygosity was significantly lower among contemporary individuals than among translocated jays ($\overline{F}_{translocated} = 0.003$, SD = 0.003, Mann-Whitney U = 355, two-tailed test, $p = 1.040 \times 10^{-4}$; $\overline{H}_{translocated} = 8.85 \times 10^{-4}$, SD = 1.57×10^{-4}, $t(16.23) = -5.791$, two-tailed test, $p = 2.61 \times 10^{-5}$), while levels of inbreeding and heterozygosity were the same between resident and translocated individuals (F, Mann-Whitney U = 111, two-tailed test, $p = 0.555$; H, $t(24.49) = 0.578$, two-tailed test, $p = 0.568$), which suggests that higher inbreeding and lower heterozygosity in the contemporary population is not accounted for by ancestral levels of inbreeding and genetic diversity alone and rather is predominantly the result of the unusually high reproductive skew among founders. These results were consistent with population-level estimates of nucleotide diversity (Table S2) and estimates of inbreeding measured as the proportion of the genome in runs of homozygosity, F_{ROH}, which were higher than F but still followed the same trends (Figure S3A). Last, our estimates of inbreeding and genetic diversity were robust to variation in sequencing depth and other experimental and analytical technicalities (Figures S3 and S4).

Reproductive skew was likely reduced by translocations but remains high

An important question surrounding the efficacy of translocations for conservation is whether the observed disparity in reproductive success among ancestors following translocations is an expected outcome from chance alone or whether the reproductive skew that we observed was unusually high. Population simulations in which the frequency of breeding pair formation, number of offspring per breeding pair, and survival varied stochastically within and across years congruent with FSJ biology showed that translocations into the resident population produced larger population sizes and significantly reduced the variance in ancestral genetic contributions compared with scenarios without translocations, even when the nonaugmented resident population exhibited growth (Mann-Whitney

$U = 655,247 - 906,343$, Bonferroni-adjusted two-tailed $p = 5.1 \times 10^{-216} - 5.1 \times 10^{-34}$ for years 2008–2022, Figure S1A). Although the observed population sizes tracked the simulated population sizes for the scenario involving translocations (Figure S1A), the observed variance and skew of ancestral genetic contributions were at or above simulated 0.95 quantiles for years 2015–2019 as well as 2009 for the variance and 2014 for skew (Figure S1B), corresponding to when the RLS-K and WSA-K lineages contributed most highly to the population. This suggests that the observed reproductive variance was higher than expected from demographic stochasticity. Importantly, when we fixed the probabilities of both pairing between adults and reproductive output among pairs to be uniform in simulations where the population demography was constrained to closely match the observed pedigree, high variance of ancestral genetic contributions was still possible (Figures S2A–S2D). However, the observed contributions of RLS-K, WSA-K, and the top 16 ancestors tended to be higher than simulated levels (Figures S2A–S2D).

Detecting reproductive skew using genetic data

We investigated approaches for characterizing reproductive skew solely from genetic data in order to (1) corroborate skew identified from the pedigree and (2) leverage the rare opportunity of having a real-life scenario involving a nearly complete population pedigree and genomic data to validate ways of identifying reproductive skew in the absence of a pedigree, which is the case for many population monitoring scenarios. We quantified the genetic contribution of sequenced ancestors to the contemporary population based on their average rank-weighted relatedness to all sequenced individuals from 2021, denoted K.

The K statistic explained a significant amount of the variation in the pedigree-based expected genetic contributions by sequenced ancestors ($r^2 = 0.62$, $t(37) = 7.7536$, two-tailed test, $p = 2.872 \times 10^{-9}$, K effect [β_K] 95% confidence interval [CI] = $[15.67, 46.44]$; see also Figure S2E). Accordingly, the distribution of K (Figure 4A) is skewed toward the same individuals with disproportionately high genetic contributions in the pedigree, and, in agreement with the pedigree, WSA-K stands out as the topmost genetic contributor (her highly successful mate, PLS-K, was not sequenced). Half of the top 10 highest genetically contributing founders to the 2021 cohort based on the pedigree are also among the top 10 individuals with the highest K (Figure 4B), while the remaining half were not sequenced, suggesting high power to identify relatively prolific lineages from genetic data alone.

DISCUSSION

We documented a 10-fold increase of the M4 CR FSJ population size within only 20 years of receiving translocated jays from neighboring subpopulations. This growth was accompanied by a geographic population expansion as new family groups established *de novo* territories throughout areas of recently restored habitat, clearly demonstrating the vital role that habitat management played in promoting population growth. While it is tempting to assume that such a demographic increase might alleviate inbreeding and temper the loss of genetic diversity, and perhaps even increase diversity when fueled by translocations, our genomic results show that this is not necessarily the case. We found that exceptionally high reproductive skew relative to levels expected from demographic

stochasticity, especially toward translocated individuals, and one breeding pair in particular that dominated contributions to the CR population for 15 years (2008–2022), curtailed the ability of translocations to decrease inbreeding and prevent genetic decay.

Generally speaking, reproductive skew reduces the effective population size (N_e), which accelerates the loss of genetic variation through drift and increases the chances of inbreeding, all of which can potentially lead to inbreeding depression and even population collapse. A stark example of this is the wolves of Isle Royale,[4–6] where over half of the population's ancestry was expected to trace back to a single male.[7] The partial genomic sweep of the combined RLS-K and WSA-K lineages throughout the CR is 40% as strong as the Isle Royale male. Although even low levels of inbreeding (pedigree $F < 4\%$) in FSJs have been shown to lower rates of survival to yearling and breeding stages,[8] the absolute increases in inbreeding that we observed are minor and unlikely to have fitness consequences at this point, though the accompanying loss of genetic diversity could reduce adaptive potential. Reproductive skew, a common feature in many natural populations spanning diverse taxa,[9,10] may have especially dire consequences within recently small, fragmented populations experiencing decreased immigration.[11] Increased inbreeding, loss of genetic diversity, and/or elevated extinction risk from reproductive skew have been documented in other cases involving birds and other taxa,[12–15] suggesting that it merits close monitoring when populations are augmented with or established via translocations. Our simulations revealed that even when mate pairing and reproductive output among breeding pairs are uniformly random, a highly optimistic scenario in terms of minimizing reproductive variance that is often violated in nature, high variability in genetic contributions can arise by chance despite population growth.

In many translocation scenarios, it will be infeasible to carry out the type of population-level monitoring required to construct the detailed pedigrees needed to characterize reproductive skew. Fortunately, we demonstrate that reproductive skew can be reliably detected and quantified from molecular data alone using the average rank-weighted relatedness statistic, K, which showed strong congruence with pedigree-based inference in terms of which lineages contributed disproportionately to the CR population. Pedigrees can also be inferred from genomic data; however, one appealing feature of the K statistic is that it can be used to estimate reproductive skew from smaller sample sizes than are needed for accurate pedigree reconstruction. Identifying reproductive skew using K is possible whenever population genetic data from at least two time points permits calculating relatedness between ancestors and descendants. Importantly, N_e can also be directly estimated from population samples sequenced at two or more time points using the temporal method.[16–18] Combining estimates of N_e with inferences of reproductive skew using K could be particularly useful for explaining small N_e to population census size ratios when demographic data are not available for characterizing reproductive variance. Such understanding of the mechanisms suppressing genetic diversity despite large population census size is extremely valuable from both an evolutionary and conservation perspective.

The ability to identify exact ancestors driving reproductive skew using K depends on their inclusion in the genetic sample (they will always have the highest K); however, if those

individuals are absent in the sample, family- or lineage-level resolution will still be possible through the proxy of sampled individuals closely related to the responsible breeder. We sequenced 57% of the CR population's founding ancestors and achieved accurate individual-level resolution into reproductive skew. In general, for characterizing reproductive skew using K, we recommend genetic sampling from as many ancestral resident or translocated individuals as possible and a representative population sample of potentially minimally related descendants, which aligns well with conventional sampling schemes for population genetic monitoring. And while K is a simple heuristic for characterizing reproductive skew that works well even at low sequencing depths, we note that with enough temporal sampling and sequencing effort to phase genomes, genomic sweeps can be characterized at haplotype-level resolution.[19]

Translocations aimed at mitigating detrimental impacts on individual organisms resulting from human land use change should still provide conservation benefits at the population, species, or ecosystem level[20]; however, their use is controversial.[21,22] Yet in the face of slated (or already completed) habitat alteration, relocating individuals may be the only option to preserve lineages and ultimately conserve populations, making the insights of our study into the mechanisms that shape the strengths and weaknesses of translocations for conservation invaluable. Despite the observed genetic decay in the CR population, we emphasize the importance of its demographic recovery, which was likely jointly fueled by translocations and habitat restoration. Larger census size provides a critical buffer against demographic and environmental stochasticity and increases the odds of population persistence.[23] Furthermore, simulations provided evidence that reproductive variance from demographic stochasticity would likely have been higher without translocations, and so the genetic state of the population would have been worse, even if the population had grown. Thus, based on our findings, we believe that translocations, even when used for mitigation, can be an effective conservation tool, while also strongly advocating the need for continued monitoring and potential management even after demographic growth following translocations. Specifically, occasional translocations of a few individuals (up to 10% of the breeding population) from other, ideally large, populations mimicking a pulse of natural gene flow in accordance with levels of migration in connected populations may be necessary for genetic rescue to offset the consequences of reproductive skew in isolated populations.

RESOURCE AVAILABILITY

Lead contact

Additional information and requests for resources should be directed to and will be fulfilled by the lead contact, Tyler Linderoth (tlindero1@msu.edu).

Materials availability

This study did not generate new, unique reagents.

Data and code availability

- Whole-genome sequence data are available from the Sequence Read Archive under BioProject accession SRA: PRJNA1099469. Demographic monitoring,

simulation, and all other data used in this study are available from Dryad: https://doi.org/10.5061/dryad.zs12jm6j0.

- Computational tools developed in this study to calculate expected genetic contributions from pedigrees and the *K* statistic are implemented in the relateStats program, which is available from Zenodo: https://doi.org/10.5281/zenodo.14260078. All other scripts and code used to perform analyses and generate figures are available at Zenodo: https://doi.org/10.5281/zenodo.14606489.
- Any additional information required to reanalyze data in this paper is available from the lead contact upon request.

STAR★METHODS

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

For our study, we monitored and sequenced Florida Scrub-Jays (*Aphelocoma coerulescens*, abbreviated FSJ) to gain high-resolution insight into the demographic and population genomic outcomes of species translocations. The FSJ is a nonmigratory, endemic bird species to Florida, USA, that experienced massive population decline starting in the twentieth century, mainly caused by degradation and loss of xeric oak scrub habitat on which it is reliant.[34,35] In 1987 FSJs were assigned Threatened status by the US Fish and Wildlife Service (USFWS). From 2003–2010, as part of a FSJ management plan developed between IMC Phosphates Company (now Mosaic Fertilizer, LLC) and the USFWS, all FSJs from five Metapopulation 4 (M4) subpopulations (51 individuals total) inhabiting areas to be mined were translocated to the M4 Core Region's Mosaic Wellfield site (Figure 1), which contained a resident population of 13 individuals.

We banded the adult FSJs in the M4 beginning in 1999 and classified any individuals banded as adults thereafter as immigrants. Starting in 2002, we conducted CR population censuses every spring before the breeding season and every July at the end of the nesting season. We identified breeding pairs and tracked their reproductive efforts through banding and field behavior observations, and from 2004 onwards monitored the nests of each family group to determine clutch size, hatching success, and fledging success. Florida Scrub-Jays are mostly genetically monogamous regardless of social factors and environment,[36] allowing for accurate assignment of parentage from behavioral observations.[37] Thus, there is strong concordance between pedigrees constructed from genetic and observational data in this species. Through 2017 young were banded as independents primarily between June and September each year. Prior to 2018, we banded and collected blood from adults after capturing them with peanut-baited walk-in traps, and then from 2018 through 2021 we banded and collected blood samples from 12-day-old nestlings. We collected blood from translocated jays at the time of capture. Blood sampling entailed first rubbing the underwing with an alcohol swab and using a needle (25 gauge for nestlings and non-translocated adults captured before 2018 and 27.5 gauge for translocated jays) to lightly puncture the brachial vein. After removing the needle, capillary tubes were used to collect the blood sample. Each sample was immediately transferred to a vial of lysis buffer labeled with the individual's

federal band number and the date of collection and stored at −80 °C for later genetic analysis.

We classified 23 jays in the population pedigree that were representative of the CR population prior to translocations as residents; comprising jays that resided in the CR prior to translocations and their direct offspring born before the end of the translocation period. We sequenced DNA from the blood of 87 individuals, encompassing 16 residents from all four family groups present in the MW prior to translocations and 28 individuals born in the CR in 2021, 18 years after translocations began. We assigned five individuals (K-WS, S-GK, PK-SL, RK-SB, WK-SL) sampled in the MW in 2004, 2006, 2007, and 2008 that were banded as adults to the resident CR group based on clear genetic affiliation to the resident population. One individual that was initially classified as a resident in the monitoring data, K-SRB, turned out to be the 2008 offspring of a male translocated from Site 12 and a CR resident mother and so was excluded from the resident group in genetic statistical comparisons so that their admixed ancestry did not bias results. We sampled 28 of the 51 translocated jays in 2003–2008, as well as one individual (K-SBP) from Site 13 that naturally immigrated to the CR between July 2009 and March 2010 which we treated as a translocated individual in subsequent pedigree and genetic analyses. In 2005, we also sampled one non-translocated individual (RSW-K) from Site 1 who migrated to Little Manatee State Park, which is outside of the CR, and so was not considered translocated nor part of the CR population in any analyses. The remaining 13 individuals were sampled in 2003–2005 from five non-donor M4 subpopulations north of the CR (Figure 1).

METHOD DETAILS

Genomic sequencing: We extracted DNA from blood samples using Qiagen DNeasy Blood and Tissue kits per the manufacturer's protocols. Whole genomes for each sample were sequenced by Novogene Corporation Inc. on an Illumina NovaSeq 6000 using 150 bp paired end sequencing. Novogene carried out basic quality control on the raw sequencing reads. Reads containing adapter sequences, greater than 10% undetermined bases ("N"), and/or more than 50% bases with Phred-scaled quality score below six were filtered out. After quality control the amount of data per individual ranged from ~77.6–148.9 million reads (median = 12.6 Gbp).

Genetic data quality control and mapping: Prior to mapping we performed additional quality control on the sequencing reads. First, we combined reads produced across different sequencing lanes for the same individual and removed reads likely derived from PCR and optical duplicates using Super Deduper[24] version 1.3.3. We then used cutadapt[25] version 3.7 to (1) trim residual adapters if at least the final three bases on the 3' end of reads aligned to the start of the adapter sequences, (2) trim missing bases ("N") from the ends of reads, and (3) remove pairs of reads for which at least one read was shorter than 70 bp after trimming. The clean reads for each individual were then separated according to sequencing lane so that the lane identity for different sequencing batches could be recorded in the alignment files.

We mapped quality-controlled reads to a high-quality draft version of the FSJ genome assembly generated by Romero et al.[38] using bwa mem[26] version 0.7.17 under default

parameters. We used the 'merge' utility of SAMtools[27] version 1.15.1 to combine aligned reads from different sequencing lanes into a single alignment (BAM) file per individual. SAMtools was used to add mate score tags to the reads with 'fixmate', sort the BAM files, and identify and record any residual duplicate reads using 'markdups'. We trimmed overlapping sequence from the mate of each read pair with the lower average quality score using the clipOverlap function of BamUtil[28] version 1.0.15. The median and mean sequencing depth per sample after quality control and mapping was 6.9x and 7.2x (SD = 1.0), respectively.

Sex assignment from genomic sequencing data: In order to assign sex to individuals without preexisting sex information (n = 53), we used SAMtools 'coverage' to calculate average autosomal depth and the average Z chromosome depth for uniquely mapped reads not sourced from PCR or optical duplicates (SAMtools arguments --ff UNMAP,SECONDARY,QCFAIL,DUP -d 0). Individuals with a Z chromosome to autosome average depth ratio below 0.6 were classified as females and individuals for which this ratio was greater than 0.6 as males since we expect the Z-chromosome depth to be approximately half that of the autosomes for (heterogametic) females. Sex assigned in this manner agreed with all 34 cases where sex had been assigned in the field based on behavioral observations or from an amplicon-based PCR assay.

Genetic variant identification: We identified genetic variants using BCFtools[27] version 1.15.1 by first calculating individual genotype likelihoods (GLs) with the mpileup function, considering only uniquely mapped, non-duplicate reads with a minimum Phred-scaled mapping quality of 20 and base quality of 13. We required at least two reads or 5% of the total reads be gapped within an individual for considering candidate indels. All other parameters were set to defaults for calculating genotype likelihoods. We called variants jointly across all 87 individuals using the BCFtools multiallelic 'call -m' model with the heterozygosity prior (-P) set to 0.003. Diploidy was assumed throughout except for Chromosome 24 (Z chromosome) for which we set females to haploid. The identity of W chromosome scaffolds was less resolved than for the Z chromosome at the time of variant calling, therefore we did not specify W scaffolds to BCFtools 'call' in order to avoid biasing calling in regions falsely classified as W-linked, noting that males will be called homozygous at W-linked sites unless there are genotyping errors.

We masked genomic sites deemed unreliable for downstream inference based on the distributions of sequencing and mapping quality metrics calculated from all reads at autosomal (excluding the Z and W chromosomes) and Z chromosome sites separately. Specifically, we masked out any sites with total depth (depth summed across all individuals) below the 0.5 quantile (523x for autosomal sites, 386x for Z chromosome sites) or total site depth above the 0.95 quantile (741x for autosomes, 681x for Z chromosome). We also masked sites meeting at least one of the following criteria: (1) fraction of reads with map quality of zero greater than 2× the genome-wide average (10%), (2) root mean square (RMS) map quality below 35, (3) fraction of reads with base quality of zero greater than 2× the genome-wide average (0.3%), (4) RMS base quality below 20. We also filtered sites using quality metric distributions based on the subset of high-quality reads, defined

as those having base and map qualities of at least 20 and 13, respectively, used for variant identification. We masked sites with total depth (VCF DP flag) above or below the respective autosomal/Z chromosome-wide median site depths +/− 25% (< 474x or > 790x for autosomes, < 336x or > 560x for Z chromosome) and/or with an exact test *p-value* for excess heterozygosity below the 2nd percentile of the autosomal/Z chromosome-wide distributions (0.06 for both autosomes and the Z chromosome). We filtered for minimum individual genetic representation from different populations or groups by requiring that at least 83% of individuals from each of the CR resident, CR contemporary, CR jays banded as adults, and translocated groups and the collective set of other M4 subpopulations were covered by at least three high quality reads and have called genotypes in order to retain a site for downstream analyses (noting that genotype calls were only used to filter missing data and not used in any analyses). We also required that at least 83% of individuals in each of the five aforementioned groups have genotype quality of at least 15 at a potentially variable site, otherwise it was masked out.

All downstream population genetic analyses were limited to sites from 31 scaffolds homologous to zebra finch autosomes (we excluded the Z and W chromosomes). We confirmed that the Z and W chromosomes were excluded from all downstream analyses based on genome alignments using Minimap2[29] between the version of the genome assembly used in our study and the final version of the assembly[38] for which the Z and W scaffolds were confidently identified based on sequencing depth and homology to zebra finch sex chromosomes.

Simulations of expected genetic contributions: We performed two types of simulations to evaluate levels of expected ancestral genetic contributions when pairing between mature, unpaired individuals and the number of offspring produced by each breeding pair is random. For the first type of simulation we constrained the pedigree structure to match the number of breeding pairs (conditional on there being available individuals to pair), the total number of offspring produced each year (given enough nesting pairs existed), and age structure of the observed pedigree in order to very closely match the observed population growth. For these simulations, the probability among unpaired adults for forming pairs and the number of offspring among pairs each year were uniformly distributed. We used these simulations to approximate lower bounds for the maximum reproductive variance due to chance given the observed demographic scenario, which provided insight into how concerned we generally need to be about reproductive variance even in the most optimistic scenario of uniformly random mate pairing and reproductive output among pairs. For the second type of simulation we sought to characterize the distribution of variance in ancestral genetic contributions by allowing the frequency of breeding pair formation, annual reproductive output (number of offspring) among breeding pairs, and survival to vary stochastically using parameter distributions relevant for FSJs, while maintaining that all unpaired adults had equal probability of pairing. These simulations not only provided more realistic distributions for reproductive skew in FSJs, but also allowed us to explore how different demographic scenarios (e.g., cases with or without translocations) affected variance in ancestral genetic contributions since the models were no longer strictly constrained by the pedigree, rather only parameterized by it.

We compared three demographic scenarios: (1) A resident population that did not receive translocations and which did not exceed four breeding pairs at any given time (which keeps the population size relatively constant and reflects the state of the actual CR population prior to translocations), (2) a resident population that did not receive translocations and with no constraints on the number of breeding pairs until a carrying capacity was reached if the population grew, (3) a resident population that received translocations and with no constraints on the number of breeding pairs until it reached the pair carrying capacity. We describe the simulation methodologies in detail below.

For the first type of simulation with demography constrained to closely match the observed pedigree and uniform mate pairing and pair reproductive output probabilities, each simulation started at time zero, t_0, and proceeded forward in time, with each time step being one year. The first event of each year introduced new individuals into the population that did not enter through local reproduction. Each time an individual entered the observed pedigree either as a resident, translocated individual, or natural immigrant, a corresponding individual was introduced into the simulated population at the analogous time step, acquiring the same lifespan as their real-life counterpart. All individuals with missing sex information were randomly assigned as either male or female with equal probability. Subsequent events occurred in the following order each year:

1. Individuals were removed from the population if they died or emigration occurred in the observed pedigree at the corresponding time step. Simulated emigrants were chosen at random among individuals that were at least one year old and not in a breeding pair (since such FSJ emigrants would be unlikely otherwise). When an individual in a breeding pair died, their mate was added back into a pool of adults available for pairing.

2. If the number of simulated breeding pairs was less than the observed number at the corresponding time step, breeding pairs in the focal population were formed by randomly pairing males and females from the pool of unpaired adults (so long as this pool contained a nonzero number of males and females), thus modeling approximately the same number of nesting pairs over time as what we observed. An individual was considered an adult if they were at least two years old (few FSJs breed before this age). The only way a breeding pair dissolved is if one or both mates died, which is in accordance with the monogamous mating system of FSJs. A single external pool of emigrants was managed in the same way to allow for reproduction outside of the focal (CR) population, enabling us to track the ancestry of immigrants which we knew were born to focal population emigrants based on the observed pedigree.

3. Locally born offspring entering the observed pedigree at the analogous time step were randomly assigned parents among the set of breeding pairs with uniform probability, modeling reproduction. Each breeding pair was not allowed to produce more than five offspring in a year since this would be rare for FSJs. Therefore, the maximum number of offspring entering the simulated population at a given time step was $\min(n_{LR}, 5n_{bp})$, where n_{LR} is the number of locally born offspring entering the observed pedigree at the corresponding time and

n_{bp} is the simulated number of breeding pairs that existed in the given time step. Offspring assigned parents in the simulation maintained the lifespan of their real-life counterparts. Reproduction was carried out in the same way for emigrants.

This process was iterated 21 times, reflecting 21 years spanned by the observed pedigree, producing a simulated pedigree for a scenario where mate pairing and reproduction were random with structure in terms of the number of breeding pairs and population sizes over time that approximately matched the observed pedigree. The demographic match is approximate because in the simulations individuals were only allowed to pair if they were at least two years old, whereas occasionally FSJs will pair when they are younger than this (less than 10% of pairs in the CR involved individuals younger than two years old) leading to rare mismatches between the number of pairs in the observed pedigree and the number of pairs in the simulations, which could potentially result in a slightly smaller simulated population size. These discrepancies in pair numbers and population sizes are small such that the simulated populations still closely matched the growth of the observed population. This pedigree-based simulation procedure, which models important biological features of FSJs such as monogamy and age structure from generational overlap, is implemented in the simPed.R script at https://github.com/tplinderoth/M4_FSJ_translocations/blob/main/scripts/simPed.R. We simulated 10k pedigrees, which we used to calculate expected ancestral genetic contributions with relateStats (see "pedigree analysis" STAR Methods section).

For the second type of simulations that modeled more realistic levels of demographic stochasticity for FSJs, the following events occurred at each time step starting at t_0 (analogous to the year 2002 in the observed pedigree):

1. Residents, translocated individuals, and natural immigrants entering the observed pedigree at the analogous time step were inserted into the population with their observed lifespan and sex. Note that the timing and amount of natural migration and translocations (if the demographic scenario involved the latter) exactly matched the observed pedigree.
2. All deceased individuals and emigrants were removed from the population. If one mate in a breeding pair died the surviving individual was returned to the pool of adult individuals available for pairing.
3. New breeding pairs form. The number of new breeding pairs at time t, k_t, was treated as a Poisson random variable, $f(k_t; \lambda_t) = \frac{\lambda_t^{k_t} e^{-\lambda_t}}{k_t!}$. The Poisson rate parameter at time t, λ_t, was given by a generalized linear model relating the number of new nesting pairs to the number of unpaired, mature individuals in the population, m. Specifically, we fit Poisson regression models for the observed number of new pairs that formed each year to m of the corresponding year, the number of currently existing pairs in the population prior to new pairing, and interaction between these two predictors, and found that the best supported model based on comparing the residual deviance of nested models involved only the main effect of m ($D = 15.967$, p-value = 0.53), which was highly

predictive of the number of nesting pairs. For each year in the simulation, $\lambda_t = e^{0.93120 + 0.03563 m_t}$. All unpaired adults had the same probability of pairing with an unpaired adult of the opposite sex. Individuals randomly pair until k_t pairs form so long as there were unpaired, mature males and females to pair, and the number of existing pairs was below a specified threshold (e.g., pair carrying capacity), so k_t is actually the maximum number of new pairs that form at time t. Age of sexual maturity was set to 2 years.

4. Breeding pairs produce offspring. Each breeding pair produced g offspring with $P(G = g)$ set according to the empirical distribution of offspring number per pair in a breeding season calculated from the entire observed pedigree. Offspring had equal probability of being male or female. The probability that an offspring individual survived y years, $P(Y = y)$, was given by the empirical distribution of survival among offspring born into the focal population in the corresponding year of the observed pedigree in order to mirror variation in survival among years that could likely be attributable to environmental factors. If there was not more than 18 observed pedigree offspring in the corresponding year to estimate the empirical survival cumulative distribution function from, we assigned offspring survival using a Poisson mixture model: $f\left(y; \alpha, \eta, \phi\right) = \alpha\frac{\eta^{y}e^{-\eta}}{y!} + \left(1-\alpha\right)\frac{\phi^{y}e^{-\phi}}{y!}$. This provides a flexible way to model bimodal survival distributions representing high mortality before individuals reach breeding age. We fit the Poisson mixture model to the empirical distribution of survival of all offspring born in the CR before 2007, which likely avoids right censuring given the maximum age observed in FSJs (> 16 years), using the L-BFGS-B algorithm implemented in the 'optim' function of R[30] version 4.4.1. The estimated maximum likelihood model was $f(y) = 0.656\frac{0.189^{y}e^{-0.189}}{y!} + 0.344\frac{5.446^{y}e^{-5.446}}{y!}$.

We iterated this process 21 times for each simulation run. It is worth pointing out that these simulations still capture important features of FSJ biology like monogomy and generational overlap, but allow for greater, and more realistic demographic stochasticity compared to the first type of simulations. This flexible simulation approach is implemented in the simJay.R script at https://github.com/tplinderoth/M4_FSJ_translocations/blob/main/scripts/simJay.R. We performed 1000 simulations for each of the three demographic scenarios and analyzed the resulting pedigrees with relateStats. We calculated the variance and skew of the genetic contribution of ancestors to the population relative to the total contribution by all ancestors for each year in the pedigree. We quantified reproductive skew as $\frac{\kappa_3}{\sigma^3}$, where $\kappa_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3$, $\sigma^3 = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{\frac{3}{2}}$, and x_i is the genetic contribution proportion of the ith ancestor among n ancestors that could potentially contribute to the population in a given year.

QUANTIFICATION AND STATISTICAL ANALYSIS

Pedigree analysis: The expected number of genomic copies originating from each resident and translocated individual among all extant CR individuals as well as within CR cohorts was calculated annually based on a nearly complete population pedigree using the 'pedstat' function of relateStats (https://github.com/tplinderoth/PopGenomicsTools). This function implements the algorithm by Hunter et al.[40] for calculating expected genomic contributions from a pedigree and relies on an additive genetic relatedness matrix, which was inferred using the 'makeA' function of the nadiv R package.[41] The expected number of genome copies, n_i, originating from a focal ancestor i from time t in a descendant group (population or cohort) consisting of individuals $M_{t+\Delta_t}$ from a later time point, $t + \Delta_t$, is $n_i = \sum_{j=1}^{M_{t+\Delta_t}} A_{i,j}$. $A_{i,j}$ is an element of matrix **A**, which consists of relatedness coefficients, $r_{i,j}$, between i and the j th individual of $M_{t+\Delta_t}$ calculated from a pedigree for which links between individual i and their ancestors are removed. Cohorts were defined as all individuals born in the CR in a given year. Reproductive skew was inferred as disproportionately high expected genetic representation from a particular ancestral individual or group relative to the total genomic representation from all 74 founding ancestors.

Population genetic characterization: We calculated genotype likelihoods (GLs) at all high-quality, autosomal sites with ANGSD[30] version 0.937 using the SAMtools model (-GL 1), considering only reads with minimum Phred-scaled map and base qualities of 20 for downstream estimates of allele frequencies, population structure, genetic diversity, and inbreeding. We assessed spatio-temporal genetic structure using a principal component analysis (PCA) based on genotype posterior probabilities calculated from the genotype likelihoods and an estimate of the minor allele frequency (ANGSD -doMaf 1 -doMajorMinor 1) under Hardy-Weinberg equilibrium. We used the genotype posterior probabilities to estimate the genetic covariance among all individuals based on biallelic SNPs with a minimum minor allele frequency (MAF) of 2% using ngsCovar, part of the ngsTools[31] software package. We performed the PCA by decomposing the genetic covariance matrix using the 'eigen' function of R. We also quantified genetic differences between spatial and temporal groups in terms of F_{ST} by first estimating the likelihoods of all possible alternate allele frequencies at each high-quality, biallelic SNP from the genotype likelihoods for resident CR individuals, contemporary CR individuals, the collective group of all translocated jays, and the collective group of all donor population jays (this group included one non-translocated individual) separately with ANGSD -doSaf 1. We used these per SNP allele frequency likelihoods to obtain a maximum likelihood estimate of the pairwise joint allele frequency spectrum between groups using the realSFS subprogram of ANGSD. We then used the allele frequency likelihoods for each group and the prior probabilities for observing joint allele frequencies provided by their respective 2-dimensional allele frequency spectrum to estimate the within and between group genetic variance components of Reynold's F_{ST} estimator with realSFS. "Weighted" genome-wide F_{ST} were calculated as the ratio of the sum of between group to the sum of the total (*i.e.*, within plus between group) genetic variance over all SNPs.

We used ngsRelate[32] to estimate relatedness between all pairs of individuals as the proportion of homologous alleles that are identical by descent (IBD) while accounting for the fact that individuals could be inbred.[42] Specifically, the genotype likelihoods and allele frequency estimates from BCFtools at all quality-controlled SNPs were used to obtain maximum likelihood estimates for all nine condensed Jacquard coefficients that represent the possible IBD states between two diploid individuals, from which relatedness, *r*, was calculated

We estimated individual inbreeding coefficients, *F*, with ngsF[33] using population-specific allele frequencies estimated from maximally large, randomly chosen subsets of individuals for which all pairwise relatedness estimated from genetic data was below 0.4. We subsetted individuals in order to avoid biasing population allele frequency estimates with the inclusion of first degree relatives. Inbreeding was estimated for individuals using allele frequencies specific to the population to which they belonged. In the case of translocated jays, with the exception of the Site 13 subpopulation, and jays from the peripheral M4 subpopulations (Sun City, Golden Aster Scrub, West Balm, Brigman, Duette Headwaters), sample sizes were too low (< 10 individuals) to accurately estimate population-specific allele frequencies. Consequently, we used allele frequencies estimated from the pool of all donor site individuals with $r < 0.4$ to estimate *F* for individuals from Site 1, Site 12, Site 18, and Texaco. Likewise, we used the historic metapopulation allele frequencies estimated from all jays sampled from 2003–2008 with $r < 0.4$ to estimate *F* for the peripheral M4 subpopulation individuals. We estimated each individual's inbreeding coefficient using their genotype likelihoods at all high-quality sites segregating within the group to which they belonged. Pruning highly related individuals in the presence of family structure can potentially reduce the accuracy of allele frequency estimates and quantities that rely on them,[43] while estimates of inbreeding that measure deviations from Hardy-Weinberg expectations (like ngsF) can be sensitive to specification of the reference population used for estimating allele frequencies. Therefore, we also estimated *F* using population-specific allele frequencies estimated without relatedness pruning as well as with respect to the metapopulation-wide allele frequencies (meaning the same allele frequencies were used for all individuals) in order to assess how sensitive our estimates of inbreeding were to different analytical procedures (Figures S3B and S3C). For the ngsF expectation maximization algorithm we used an initial estimate of *F* calculated from the GLs, and set the maximum number of iterations (--max_iters) to 3000, the minimum number of iterations (--min_iters) to 10, and the maximum root mean square deviation between iterations (--min_epsilon) to 1×10^{-5} as the convergence criterion.

We also estimated individual inbreeding based on runs of homozygosity (ROH) identified with the hidden Markov model (HMM) implemented in BCFtools roh[44] using population-specific allele frequencies as well as metapopulation-wide allele frequencies as in the ngsF analysis (Figure S3C) and a constant per base pair recombination rate of 3.6×10^{-8}, representing a multi-avian-species average.[45–49] For each individual, we trained and ran the HMM using all high-quality sites segregating within the group to which they belonged or at all SNPs when using metapopulation allele frequencies with the convergence threshold set to 1×10^{-10} for estimating the HMM parameters with the Viterbi algorithm. We estimated

individual inbreeding, F_{ROH}, as the proportion of the genome in ROH, considering only sites for which there was less than a 1% chance that the inferred state (autozygous or non-autozygous) was an error as determined by the HMM forward-backward algorithm.

We estimated heterozygosity, H, of each individual using their genotype likelihoods to calculate the posterior probability that they had 0, 1, or 2 alternate alleles at each quality-controlled site in the genome while accounting for their level of inbreeding[33] as measured by F_{ROH} (estimated using population-specific allele frequencies, noting that BCFtools roh was robust to different reference populations) using ANGSD (-doSaf 2 supplying the reference genome to -anc and F_{ROH} to -indF). Given s analyzed sites, an individual's per site posterior expected H is given by $\frac{1}{s}\sum_{i=1}^{s} p_{1,i}$, where $p_{1,i}$ is the posterior probability that site i has one alternate allele. We also estimated population nucleotide diversity, π, from the posterior expected site frequency spectrum (SFS) for each population.[50] Specifically, we used the MAF likelihoods at each site calculated with ANGSD -doSaf 1 and prior probabilities for the MAF at any given site provided by an estimate of the population's folded genome-wide SFS that accounted for inbreeding[33] (ANGSD -doSaf 2) to obtain a posterior expected estimate of the population's SFS from which we calculated π using realSFS saf2theta.

We used random subsets of individuals for which all pairwise relatedness was below 0.4 to test for significant differences in the distribution of individual inbreeding coefficients and mean heterozygosity between contemporary, resident, and translocated groups with two-sided Mann-Whitney U tests and Welch's t-tests. We ensured that estimates of inbreeding and heterozygosity were not confounded by differences in sequencing depth by fitting linear models in R to test whether variation in average individual sequence depth explained variation in estimates of inbreeding and heterozygosity (Figures S4A–S4C). We also performed an ANOVA in R to test for differences in average depth among individuals belonging to the contemporary CR, resident CR, donor population, and other M4 groups (Figure S4D). For each linear model, we performed 10k bootstrap replicates based on sampling vectors of observations to estimate bias-corrected accelerated confidence intervals for the effect size of sequencing depth (for models of inbreeding and heterozygosity as a function of individual sequence depth) or group (for the sequence depth ANOVA).

Assessing reproductive skew from genomic data: We used the genetic estimates of relatedness to calculate the average rank-weighted relatedness, K, between all sequenced resident and translocated jays to the contemporary sample. For focal ancestor i, descendant individual j, and another randomly chosen ancestor m, $j \neq m$, $k_i = \frac{1}{n_d}\sum_{j=1}^{n_d} r_{i,j} P(r_{i,j} > r_{m,j})$, where n_d is the number of descendant individuals and $r_{i,j}$ is the probability that alleles from i and j are IBD, specifically, the relatedness coefficient between i and j. Given n_a ancestral individuals, $P(r_{i,j} > r_{m,j}) = \frac{1}{n_a - 1}\sum_{m \neq i}^{n_a} \mathbb{I}_{r_{i,j} > r_{m,j}}$. With higher relatedness to a descendant population, the K statistic for a focal ancestor will increase proportional to the probability that they are more related to descendants than other ancestors. Comparing values for K among ancestors is useful for identifying and assessing relative degrees of reproductive skew. We fit a linear model in R to determine how much of the variation in

the pedigree-based expected genetic contributions was explained by the genetic K statistic, and tested for whether this amount of accounted for variation was significant using a two-tailed t-test and bias-corrected accelerated 95% confidence intervals for the effect of K, β_K, constructed from 10k iterations of bootstrapping vectors of observations to estimate β_K.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We would like to posthumously thank Dr. Reed Bowman for his foresight and long-term (>20 years) advising on data collection and direction of FSJ translocations in the M4 metapopulation and for his positive impacts on FSJ conservation at large. We would also like to thank David Gordon, who managed the initial project, banding, and field data collections for 15 years, as well as Lee Walton and Tyler Buckley, who also collected blood samples. We appreciate members of the Fitzpatrick Lab and four reviewers for their helpful feedback and discussion, which greatly improved this study. Funding was provided by Mosaic Fertilizer, LLC, and its predecessor IMC Phosphates; National Science Foundation grant DEB 2016569 (S.W.F.); and National Institutes of Health grant R35 GM133412 (N.C.). This is Kellogg Biological Station contribution 2370.

REFERENCES

1. Berger-Tal O, Blumstein DT, and Swaisgood RR (2020). Conservation translocations: a review of common difficulties and promising directions. Anim. Conserv 23, 121–131. 10.1111/acv.12534.
2. U.S. Fish and Wildlife Service (2001). Scrub-Jay population modeling. section 2.4 appendix 3. In Biological Opinion and Associated Florida Scrub-Jay Habitat Management Plan for IMC Phosphates Company Southern Hillsborough and Manatee County Projects (U.S. Fish and Wildlife Service).
3. Bowman R. (2008). Modelling the Effect of Different Mitigation Options for Florida Scrub-Jays (Aphelocoma coerulscens) on the Four Corners/Lonesome Regional Permit Area on the Long-Term Persistence of the Scrub-Jay Metapopulation in Southern Hillsborough and Manatee Counties (IMC-Agrico Company).
4. Adams JR, Vucetich LM, Hedrick PW, Peterson RO, and Vucetich JA (2011). Genomic sweep and potential genetic rescue during limiting environmental conditions in an isolated wolf population. Proc. Biol. Sci 278, 3336–3344. 10.1098/rspb.2011.0261 [PubMed: 21450731]
5. Robinson JA, Räikkönen J, Vucetich LM, Vucetich JA, Peterson RO, Lohmueller KE, and Wayne RK (2019). Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. Sci. Adv 5, eaau0757. 10.1126/sciadv.aau0757
6. Hedrick PW, Robinson JA, Peterson RO, and Vucetich JA (2019). Genetics and extinction and the example of Isle Royale wolves. Anim. Conserv 22, 302–309. 10.1111/acv.12479.
7. Hedrick PW, Peterson RO, Vucetich LM, Adams JR, and Vucetich JA (2014). Genetic rescue in Isle Royale wolves: genetic analysis and the collapse of the population. Conserv. Genet 15, 1111–1121. 10.1007/s10592-014-0604-1
8. Chen N, Cosgrove EJ, Bowman R, Fitzpatrick JW, and Clark AG (2016). Genomic consequences of population decline in the endangered Florida Scrub-Jay. Curr. Biol 26, 2974–2979. 10.1016/j.cub.2016.08.062. [PubMed: 27746026]
9. Benvenuto C, and Lorenzi MC (2023). Social regulation of reproduction: control or signal? Trends Ecol. Evol 38, 1028–1040. 10.1016/j.tree.2023.05.003. [PubMed: 37385846]
10. Árnason E, Koskela J, Halldórsdóttir K, and Eldon B. (2023). Sweepstakes reproductive success via pervasive and recurrent selective sweeps. eLife 12, e80781. 10.7554/eLife.80781.
11. McCarthy MS, Lester JD, Cibot M, Vigilant L, and McLennan MR (2020). Atypically high reproductive skew in a small wild chimpanzee community in a human-dominated landscape. Folia Primatol. (Basel) 91, 688–696. 10.1159/000508609. [PubMed: 32604094]

12. Jamieson IG (2011). Founder effects, inbreeding, and loss of genetic diversity in four avian reintroduction programs. Conserv. Biol 25, 115–123. 10.1111/j.1523-1739.2010.01574.x. [PubMed: 20825445]
13. Kelly MJ (2001). Lineage loss in Serengeti cheetahs: Consequences of high reproductive variance and heritability of fitness on effective population size. Conserv. Biol 15, 137–147. 10.1111/j.1523-1739.2001.99033.x
14. Harvey Sky N, Jackson J, Chege G, Gaymer J, Kimiti D, Mutisya S, Nakito S, and Shultz S. (2022). Female reproductive skew exacerbates the extinction risk from poaching in the eastern black rhino. Proc. Biol. Sci 289, 20220075. 10.1098/rspb.2022.0075.
15. Stojanovic D, McLennan E, Olah G, Cobden M, Heinsohn R, Manning AD, Alves F, Hogg C, and Rayner L. (2023). Reproductive skew in a vulnerable bird favors breeders that monopolize nest cavities. Anim. Conserv 26, 675–683. 10.1111/acv.12855.
16. Krimbas CB, and Tsakas S. (1971). The genetics of Dacus oleae. V. Changes of esterase polymorphism in a natural population following insecticide control—Selection or drift? Evolution 25, 454–460. 10.1111/j.1558-5646.1971.tb01904.x. [PubMed: 28565021]
17. Nei M, and Tajima F. (1981). Genetic drift and estimation of effective population size. Genetics 98, 625–640. 10.1093/genetics/98.3.625. [PubMed: 17249104]
18. Waples RS, and Yokota M. (2007). Temporal estimates of effective population size in species with overlapping generations. Genetics 175, 219–233. 10.1534/genetics.106.065300. [PubMed: 17110487]
19. Viluma A, Flagstad Ø, Åkesson M, Wikenros C, Sand H, Wabakken P, and Ellegren H. (2022). Whole-genome resequencing of temporally stratified samples reveals substantial loss of haplotype diversity in the highly inbred Scandinavian wolf population. Genome Res. 32, 449–458. 10.1101/gr.276070.121. [PubMed: 35135873]
20. IUCN/SSC (2013). Guidelines for Reintroductions and Other Conservation Translocations. Version 1.0 (IUCN Species Survival Commission).
21. Germano JM, Field KJ, Griffiths RA, Clulow S, Foster J, Harding G, and Swaisgood RR (2015). Mitigation-driven translocations: are we moving wildlife in the right direction? Front. Ecol. Environ 13, 100–105. 10.1890/140137.
22. Bradley HS, Tomlinson S, Craig MD, Cross AT, and Bateman PW (2022). Mitigation translocation as a management tool. Conserv. Biol 36, e13667. 10.1111/cobi.13667.
23. Grimm V, Revilla E, Groeneveld J, Kramer-Schadt S, Schwager M, Tews J, Wichmann MC, and Jeltsch F. (2005). Importance of buffer mechanisms for population viability analysis. Conserv. Biol 19, 578–580. 10.1111/j.1523-1739.2005.000163.x.
24. Petersen KR, Streett DA, Gerritsen AT, Hunter SS, and Settles ML (2015). Super Deduper, fast PCR duplicate detection in fastq files. In Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics (ACM), pp. 491–492. 10.1145/2808719.2811568.
25. Martin M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, 10–12. 10.14806/ej.17.1.200.
26. Li H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at arXiv. 10.48550/arXiv.1303.3997.
27. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. (2021). Twelve years of SAMtools and BCFtools. GigaScience 10, giab008. 10.1093/gigascience/giab008.
28. Jun G, Wing MK, Abecasis GR, and Kang HM (2015). An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. Genome Res. 25, 918–925. 10.1101/gr.176552.114. [PubMed: 25883319]
29. Li H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34, 3094–3100. 10.1093/bioinformatics/bty191. [PubMed: 29750242]
30. Korneliussen TS, Albrechtsen A, and Nielsen R. (2014). ANGSD: analysis of next generation sequencing data. BMC Bioinformatics 15, 356. 10.1186/s12859-014-0356-4. [PubMed: 25420514]

31. Fumagalli M, Vieira FG, Linderoth T, and Nielsen R. (2014). ngsTools: methods for population genetics analyses from next-generation sequencing data. Bioinformatics 30, 1486–1487. 10.1093/bioinformatics/btu041. [PubMed: 24458950]
32. Hanghøj K, Moltke I, Andersen PA, Manica A, and Korneliussen TS (2019). Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. GigaScience 8, giz034. 10.1093/gigascience/giz034.
33. Vieira FG, Fumagalli M, Albrechtsen A, and Nielsen R. (2013). Estimating inbreeding coefficients from NGS data: impact on genotype calling and allele frequency estimation. Genome Res. 23, 1852–1861. 10.1101/gr.157388.113. [PubMed: 23950147]
34. Woolfenden GE, and Fitzpatrick JW (1996). Florida Scrub-Jay (Aphelocoma coerulescens). In Birds of North America, Poole A, Stettenheim P, and Gill F, eds. (Cornell Lab of Ornithology). 10.2173/bna.228.
35. Fitzpatrick JW, and Bowman R. (2016). Florida scrub-jays: Oversized territories and group defense in a fire-maintained habitat. In Cooperative Breeding in Vertebrates, Koenig WD, and Dickinson JL, eds. (Cambridge University Press), pp. 77–96. 10.1017/CBO9781107338357.006.
36. Townsend AK, Bowman R, Fitzpatrick JW, Dent M, and Lovette IJ (2011). Genetic monogamy across variable demographic landscapes in cooperatively breeding Florida scrub-jays. Behav. Ecol 22, 464–470. 10.1093/beheco/arq227.
37. Quinn JS, Woolfenden GE, Fitzpatrick JW, and White BN (1999). Multi-locus DNA fingerprinting supports genetic monogamy in Florida scrub-jays. Behav. Ecol. Sociobiol 45, 1–10. 10.1007/s002650050534.
38. Romero FG, Beaudry FEG, Hovmand Warner E, Nguyen TN, Fitzpatrick JW, and Chen N. (2024). A new high-quality genome assembly and annotation for the threatened Florida Scrub-Jay (Aphelocoma coerulescens). G3 (Bethesda) 14, jkae232. 10.1093/g3journal/jkae232.
39. R Core Team (2024). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).
40. Hunter DC, Pemberton JM, Pilkington JG, and Morrissey MB (2019). Pedigree-based estimation of reproductive value. J. Hered 110, 433–444. 10.1093/jhered/esz033. [PubMed: 31259373]
41. Wolak ME (2012). nadiv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. Methods Ecol. Evol 3, 792–796. 10.1111/j.2041-210X.2012.00213.x.
42. Hedrick PW, and Lacy RC (2015). Measuring relatedness between inbred individuals. J. Hered 106, 20–25. 10.1093/jhered/esu072. [PubMed: 25472983]
43. Waples RS, and Anderson EC (2017). Purging putative siblings from population genetic data sets: a cautionary view. Mol. Ecol 26, 1211–1224. 10.1111/mec.14022. [PubMed: 28099771]
44. Narasimhan V, Danecek P, Scally A, Xue Y, Tyler-Smith C, and Durbin R. (2016). BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. Bioinformatics 32, 1749–1751. 10.1093/bioinformatics/btw044. [PubMed: 26826718]
45. Backström N, Forstmeier W, Schielzeth H, Mellenius H, Nam K, Bolund E, Webster MT, Öst T, Schneider M, Kempenaers B, et al. (2010). The recombination landscape of the zebra finch Taeniopygia guttata genome. Genome Res. 20, 485–495. 10.1101/gr.101410.109. [PubMed: 20357052]
46. Groenen MA, Cheng HH, Bumstead N, Benkel BF, Briles WE, Burke T, Burt DW, Crittenden LB, Dodgson J, Hillel J, et al. (2000). A consensus linkage map of the chicken genome. Genome Res. 10, 137–147. 10.1101/gr.10.1.137. [PubMed: 10645958]
47. Smeds L, Mugal CF, Qvarnström A, and Ellegren H. (2016). High-resolution mapping of crossover and non-crossover recombination events by whole-genome re-sequencing of an avian pedigree. PLoS Genet. 12, e1006044. 10.1371/journal.pgen.1006044.
48. Hagen IJ, Lien S, Billing AM, Elgvin TO, Trier C, Niskanen AK, Tarka M, Slate J, Saetre GP, and Jensen H. (2020). A genome-wide linkage map for the house sparrow (Passer domesticus) provides insights into the evolutionary history of the avian genome. Mol. Ecol. Resour 20, 544–559. 10.1111/1755-0998.13134. [PubMed: 31912659]
49. Bascón-Cardozo K, Bours A, Manthey G, Durieux G, Dutheil JY, Pruisscher P, Odenthal-Hesse L, and Liedvogel M. (2024). Fine-scale map reveals highly variable recombination rates associated

with genomic features in the Eurasian blackcap. Genome Biol. Evol 16, evad233. 10.1093/gbe/evad233.

50. Korneliussen TS, Moltke I, Albrechtsen A, and Nielsen R. (2013). Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. BMC Bioinformatics 14, 289. 10.1186/1471-2105-14-289. [PubMed: 24088262]

Highlights

- Translocations of Florida Scrub-Jays into restored habitat fuel population growth
- Reproductive skew restricts the effective population size and increases inbreeding
- Genetic diversity in the recipient population decreased despite population growth
- Translocations likely curtailed higher rates of inbreeding and genetic decay

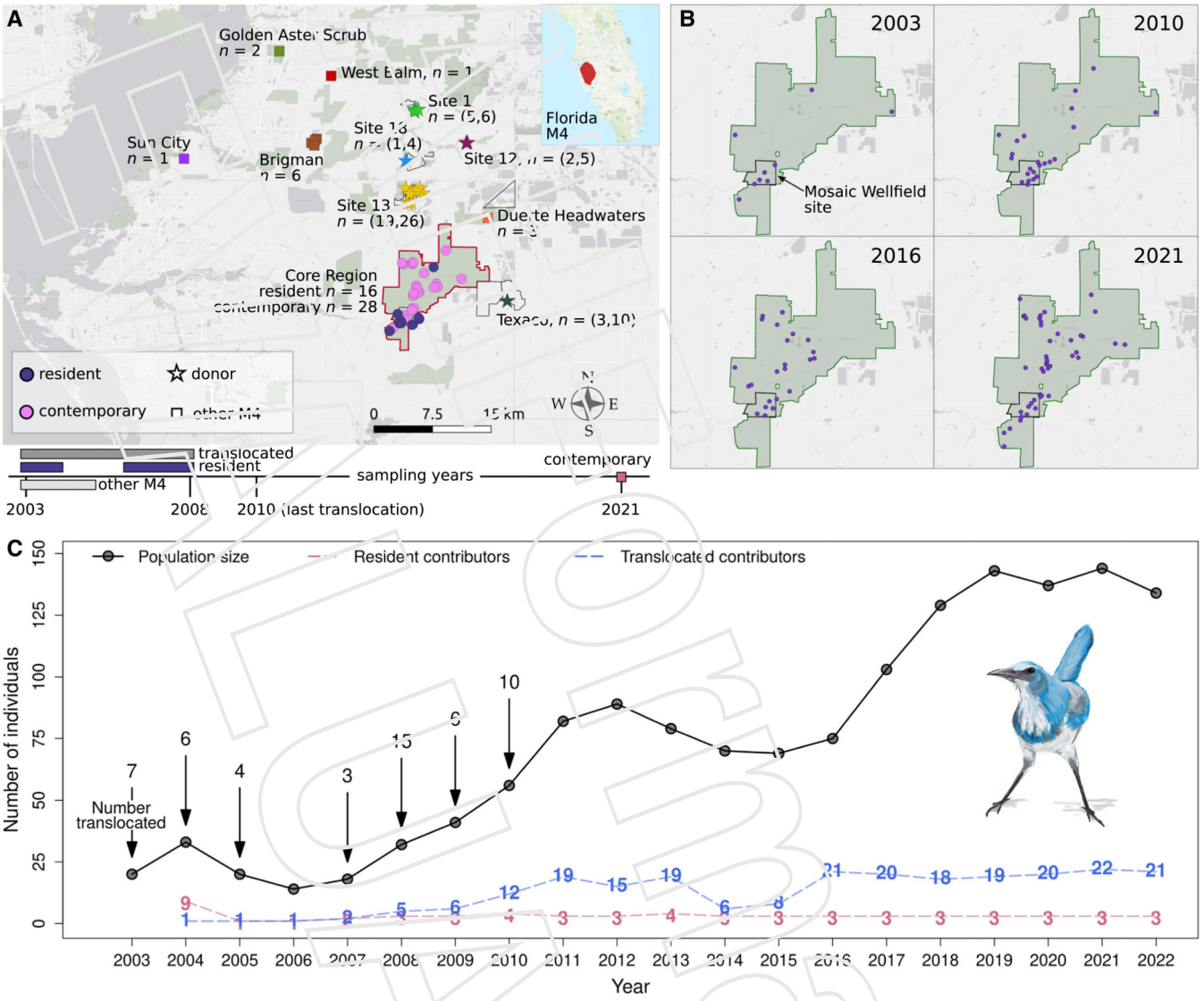


Figure 1. Demographic response of the CR FSJ population to translocations

(A) Map showing M4 subpopulation sample sizes (*n*) for genetic analyses, with individuals depicted by colored shapes. For translocation donor subpopulations, we show the genetic sample size and total number individuals translocated from the site, in that order. Individuals were translocated into the Mosaic Wellfield (MW) site and then expanded throughout the core region (CR) outlined in red. The scale bar denotes geographic distances for the main map in kilometers (km). The geographic range of the entire M4 metapopulation is colored red in the inset map of the Florida peninsula. The time periods in which blood samples were taken from individuals for genetic analyses are displayed below the map.

(B) Purple dots denoting family groups show the geographic expansion of the population throughout the CR (green area) over time. The MW site is outlined in black.

(C) The solid black line shows the census size of the CR population over time with arrows showing the year and number of jays translocated into the MW site. Dashed lines indicate

the number of translocated (blue) and resident (pink) founding lineages contributing to cohorts born in the CR over time.
See also Figure S1.

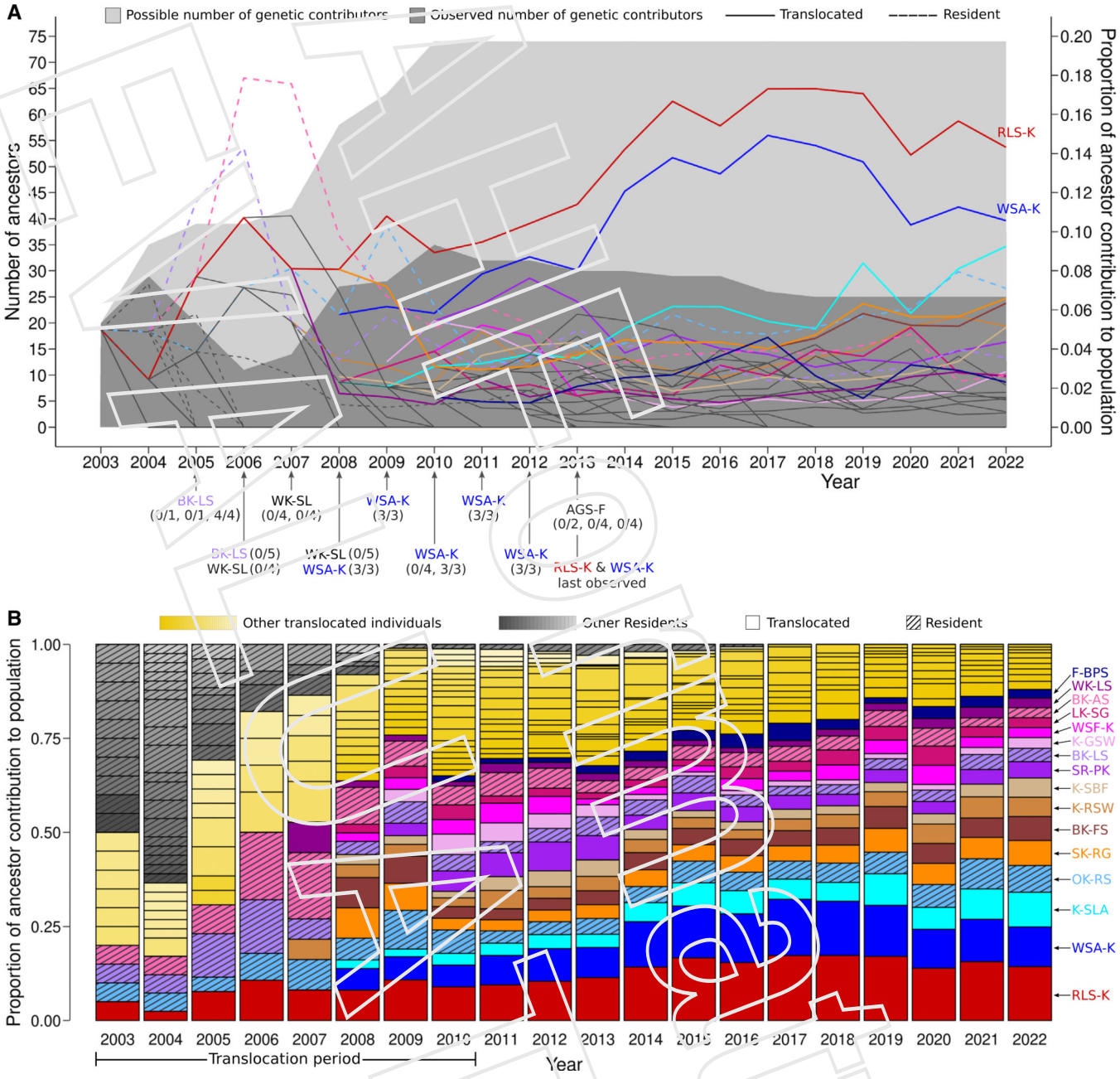


Figure 2. Expected genetic contributions by CR ancestors over time based on the population pedigree

(A) The total number of founding ancestors that had existed in the CR and could potentially contribute to the population is indicated by the top of the light gray shaded area, while the number of ancestors that were actually genetically represented in the population based on the pedigree is denoted by the top of the dark gray shaded area. The number of potential genetic contributors increased as jays were moved to the MW during the translocation period (2003–2010), with only a fraction contributing to the population over time. Lines show the expected proportion of the total founder genetic material originating from each of the 74 CR population ancestors over time. Genetic contributions were skewed toward a small subset of

CR founders, with one breeding pair comprised of RLS-K (male, red) and WSA-K (female, blue) contributing the most (see also Table S1). Annotated arrows show the mate of RLS-K each year and the number of young that the respective pair fledged out of the total number of eggs laid per nest attempt (number fledged/total number of eggs).

(B) The same expected genetic contribution by ancestors to the CR population is depicted by lines in (A) but shown as stacked bars (individuals are ordered based on ranked contribution to the 2022 population), highlighting genetic takeover by a subset of mostly translocated lineages. Sixteen founding individuals from which over 76% of the CR ancestry is expected to have derived since 2015 are labeled in (B) and denoted by the same colors in (A). The remaining ancestral lineages are colored yellow (B) or gray (lines in A; bars in B). Figures S1 and S2A–S2D show how the observed variance in ancestral genetic contributions shown here compares to levels expected from demographic stochasticity.

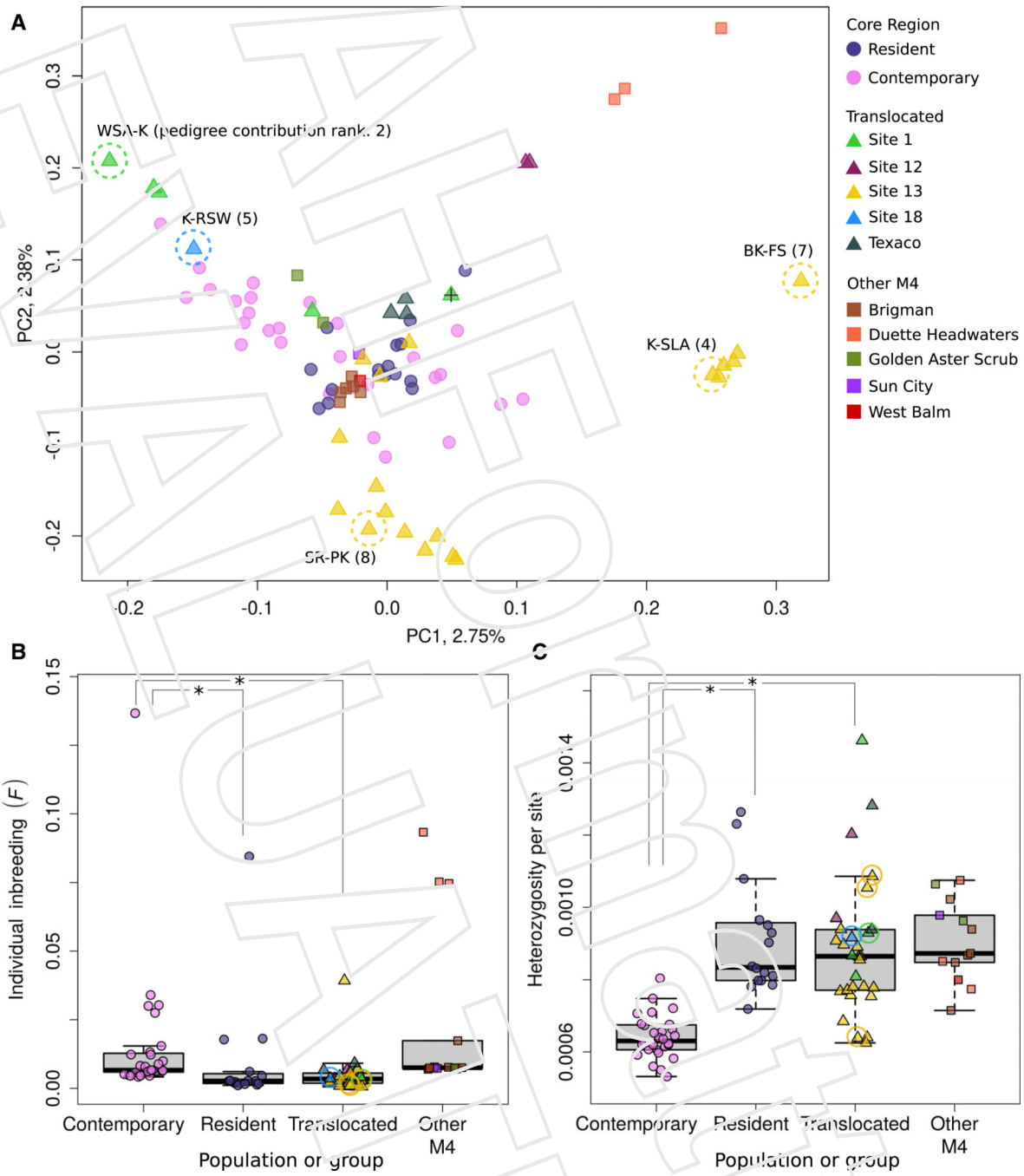


Figure 3. Spatiotemporal genomic characterization of the CR and neighboring M4 subpopulations

(A) Genetic similarity among sequenced individuals based on a principal-component analysis (PCA) of genome-wide variation in our M4 sample. The percentage of the total genetic variation explained by each PC is indicated on each axis. Sequenced founding ancestors among the top 10 largest genetic contributors to the contemporary population based on the pedigree are enclosed in dashed circles and denoted by their color-band identifiers and the rank of their pedigree-based genetic contribution to the 2022 population.

"+" denotes a donor site individual that was not translocated and not a member of the CR population.

(B and C) Comparison of the distributions of individual inbreeding coefficients (*F*) and heterozygosity (*H*) for all sequenced individuals from different groups. Colors and shapes denoting individuals are the same as in (A). Significant ($p < 0.05$) group differences in *F* and *H* from tests performed on subsets of individuals for which all pairwise relatedness was below 0.4 are indicated by an asterisk (*). No statistical tests were performed involving the "other M4" group. The same highest-ranked contributors to the contemporary CR population based on the pedigree circled in (A) are also circled in the plots of *F* and *H*. See Figure S3 for distributions of individual inbreeding estimated using different approaches and Table S2 for group-level comparisons of genetic diversity based on nucleotide diversity (π).

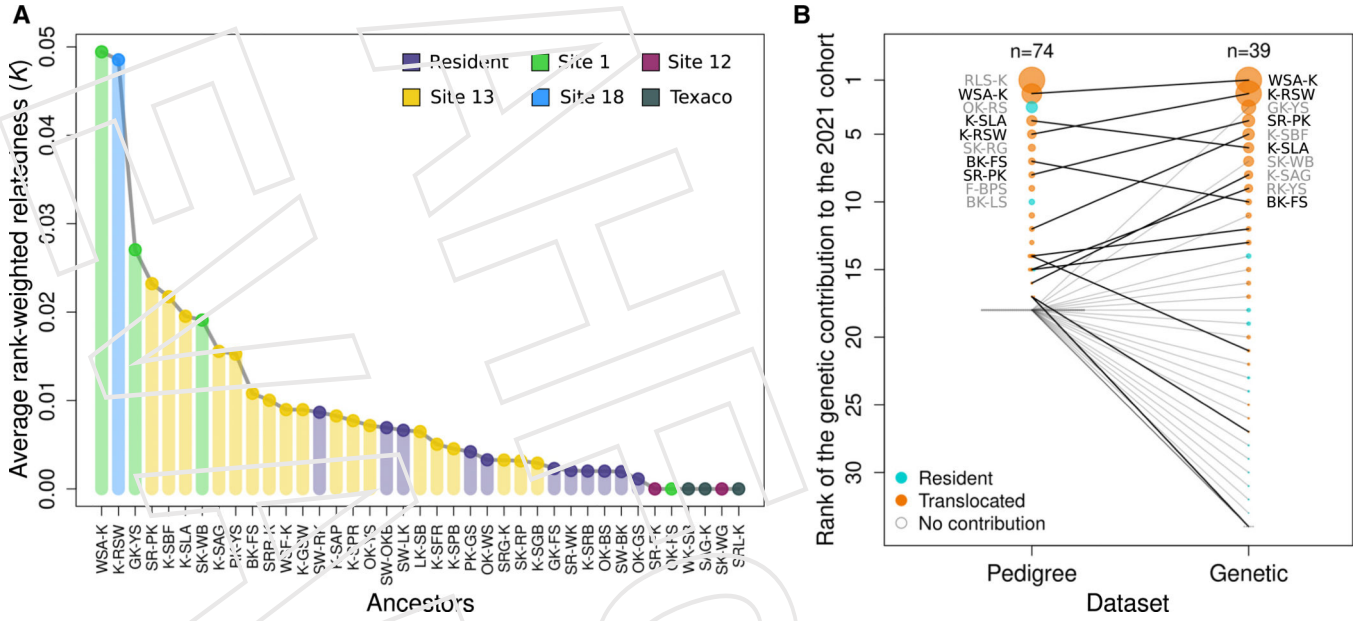


Figure 4. Characterization of reproductive skew from genomic sequencing data

(A) The distribution of the average rank-weighted relatedness statistic, *K*, for sequenced CR population founders highlights large differences in the genetic contribution by ancestors to the 2021 CR cohort.

(B) A comparison between the characterization of ancestral genetic contributions to the 2021 cohort using the pedigree with all 74 ancestors versus the genetic *K* statistic calculated for 39 sequenced ancestors shows high concordance between the two approaches for inferring reproductive skew. Resident versus translocated origin of individuals is indicated by cyan and orange colors, respectively, and the size of points is scaled to the maximum individual genetic contribution in the respective analysis. Lines connect individuals common to both analyses and are black if the individual contributed to the 2021 CR cohort according to the pedigree. The identity of the top 10 contributing jays from each analysis is displayed next to their respective points, with individuals common to both lists in black.

See also Figure S2E.

Curr Biol. Author manuscript; available in PMC 2025 May 19.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Wild Florida Scrub-Jay (*Aphelocoma coerulescens*)	This paper	N/A
Deposited data		
Florida Scrub-Jay whole genome sequencing data in fastq format	This paper	BioProject SRA: PRJNA1099469
Florida Scrub-Jay draft reference genome, M4 metapopulation genetic variants (VCF format), genomic quality masks, M4 core region Florida Scrub-Jay population pedigree, demographic monitoring data, population metadata, and population simulation data.	This paper	Dryad: 10.5061/dryad.z612jm6j0
Software and algorithms		
SuperDeduper version 1.3.3	Petersen et al.[24]	https://github.com/ibest/HTStream
Cutadapt version 3.7	Martin[25]	https://github.com/marcelm/cutadapt/
BWA-MEM version 0.7.17	Li26	https://github.com/lh3/bwa
SAMtools version 1.15.1	Danecek et al.[27]	https://github.com/samtools/samtools
BCFtools version 1.15.1	Danecek et al.[27]	https://github.com/samtools/bcftools
BamUtil version 1.0.15	Jun et al.[28]	https://github.com/statgen/bamUtil
Minimap2 version 2.28	Li29	https://github.com/lh3/minimap2
ANGSD version 0.937	Korneliussen et al.[30]	http://github.com/ANGSD/angsd
ngsTools	Fumagalli et al.[31]	https://github.com/mfumagalli/ngsTools
ngsRelate version 2	Hanghøj et al.[32]	https://github.com/ANGSD/NgsRelate
ngsF	Vieira et al.[33]	https://github.com/fgvieira/ngsF
relateStats	This paper	https://github.com/tplinderoth/PopGenomicsTools
Scripts used to simulate populations (simPed.R and simJay.R) and analyze pedigree and genetic data	This paper	https://github.com/tplinderoth/M4_FSJ_translocations