

Article

Limited domestic introgression in a final refuge of the wild pigeon



William J. Smith, Ashley T. Sendell-Price, Annette L. Fayet, ..., Steven Kelly, Lindsay A. Turnbull, Sonya M. Clegg

william.smith@queens.ox.ac.uk

Highlights

The Rock Dove is the wild ancestor of today's domestic and feral pigeons

The status of the Rock Dove was unclear owing to gene flow from its domestic relatives

We used genetic and morphological data to identify Rock Doves in the British Isles

Outer Hebridean populations experience the least gene flow with domestic pigeons

Smith et al., iScience 25, 104620
July 15, 2022 © 2022 The Author(s).
<https://doi.org/10.1016/j.isci.2022.104620>

Article

Limited domestic introgression in a final refuge of the wild pigeon

William J. Smith,^{1,8,*} Ashley T. Sendell-Price,^{1,2} Annette L. Fayet,^{3,4} Teia M. Schweizer,⁵ Michał T. Jezierski,¹ Charles van de Kerkhof,⁶ Ben C. Sheldon,¹ Kristen C. Ruegg,⁵ Steven Kelly,⁷ Lindsay A. Turnbull,⁷ and Sonya M. Clegg¹

SUMMARY

Domesticated animals have been culturally and economically important throughout history. Many of their ancestral lineages are extinct or genetically endangered following hybridization with domesticated relatives. Consequently, they have been understudied compared to the ancestral lineages of domestic plants. The domestic pigeon *Columba livia*, which was pivotal in Darwin's studies, has maintained outsized cultural significance. Its role as a model organism spans the fields of behavior, genetics, and evolution. Domestic pigeons have hybridized with their progenitor, the Rock Dove, rendering the latter of dubious genetic status. Here, we use genomic and morphological data from the putative Rock Doves of the British Isles to identify relictual undomesticated populations. We reveal that Outer Hebridean Rock Doves have experienced minimal levels of introgression. Our results outline the contemporary status of these wild pigeons, highlighting the role of hybridization in the homogenization of genetic lineages.

INTRODUCTION

The genetic status of the undomesticated Rock Dove is ambiguous throughout its original Afro-Eurasian range (Baldaccini, 2020; Stringham et al., 2012). The “wildtype” plumage of the Rock Dove regularly manifests in feral domestic pigeons, hindering the discrimination of the two forms at an individual level (Goodwin and Gillmor, 1970; Johnston et al., 1988). Subtle morphological differences have been proposed but not rigorously assessed, particularly with respect to the effects of hybridization (Johnston et al., 1988; Johnston and Janiga, 1995). In many regions where Rock Doves were once present (e.g., England, Wales, Norway, and the Italian peninsula) they have long been fully replaced by feral pigeons (Baldaccini, 2020; Brown and Grice, 2005; Lovegrove et al., 1994; Michaelsen and Refvik, 2003). In previous European strongholds (e.g., Sardinia and the Faroe Islands) observations of plumage strongly suggest that wild-domestic hybridization has been occurring for decades (Baldaccini, 2020; Fr et al., 1949; Johnston, 1992). African and Asian populations are understudied, and the status of any of their Rock Doves is, therefore, unknown (Johnston and Janiga, 1995). Although the Macaronesian (*C. l. atlantis*) and Central Asian (*C. l. nigricans*) subspecies are now considered to be of feral provenance, for those thought not to have a history of domestication (e.g., the central Egyptian *C. l. dakhlae* and west African *C. l. gymnocyclus*), the extent to which they have hybridized with feral pigeons is unclear. Given the prolonged history of domestic populations, and the porous geographical and reproductive boundaries between the two forms, there are doubts as to whether any unadmixed Rock Dove populations remain. This is a concern of particular importance with regard to the ancestral type (as the primary progenitor of the domestic pigeon) *C. l. livia* (Baldaccini, 2020; Stringham et al., 2012). Nevertheless, there is likely to be variation in the level of feral hybridization among different Rock Dove populations, and those that have experienced minimal feral introgression are of scientific and conservation interest. The caves and cliffs of certain regions of the British Isles, including the Hebrides, Shetland, Orkney, parts of the Scottish Highlands, and parts of western Ireland (e.g. Cape Clear Island), have been proposed to harbor ancestral-type Rock Dove populations that are minimally impacted by hybridization, based on the predominance of “wildtype” plumage traits (Johnston, 1992). However, feral and free-flying domestic pigeon populations frequent these regions, and Rock Dove colonies in the Highlands and Northern Isles show “feral” plumage traits (Hewson, 1967). Using a combination of whole-genome sequencing data and morphometric assessments, we first explore the contemporary status of these putative Rock Doves; and then examine patterns of domestic introgression

¹Edward Grey Institute of Field Ornithology, Department of Zoology, University of Oxford, Oxford, UK

²Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

³Oxford Navigation Group, Department of Zoology, University of Oxford, Oxford, UK

⁴Norwegian Institute for Nature Research, Trondheim, Norway

⁵Biology Department, Colorado State University, Fort Collins, CO, USA

⁶Fjällgås, Lund, Sweden

⁷Department of Plant Sciences, University of Oxford, Oxford, UK

⁸Lead contact

*Correspondence: william.smith@queens.ox.ac.uk

<https://doi.org/10.1016/j.isci.2022.104620>



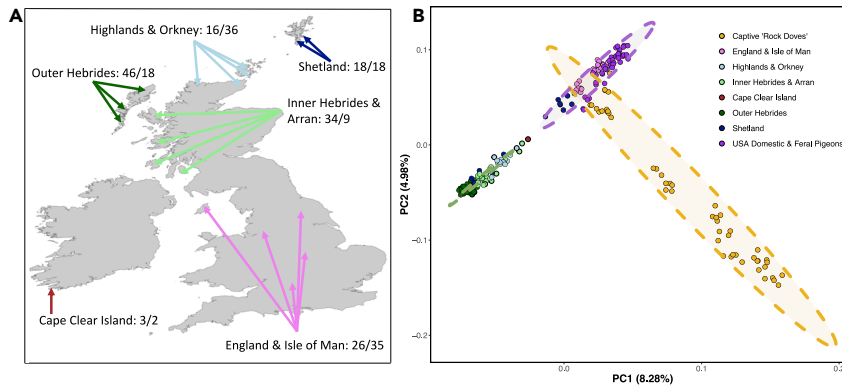


Figure 1. Rock Dove populations

(A) Rock Doves and feral pigeons were sampled from various locations around the British Isles. Rock Dove capture sites were categorized according to geography. Numbers indicate: birds sampled for DNA/adults measured for morphological data. Expanded location and sampling information is provided in [Table S1](#).

(B) Principal component analysis shows genetic differentiation of wild Rock Doves from both domestic/feral birds and captive "Rock Doves." The analysis was carried out using PCAnsd, based on a genotype likelihood file filtered for linkage disequilibrium with 2,583,744 polymorphic nuclear sites, and 232 individual birds. The variance explained by PC1 and PC2 is shown in parentheses on the axes. 95% confidence interval ellipses show the wild Rock Dove (green), feral/domestic pigeon (purple), and captive "Rock Dove" (amber) groups. The size and shape of the Rock Dove ellipse reflect the close genetic relationships between different populations within the British Isles. Color of points represents the population grouping to which the individual bird belongs (see legend).

among populations, among individuals, and along the genome, to assess the value of Rock Dove populations in the British Isles as one of the last bastions of ancestral variation in this species.

RESULTS

We characterized genetic structure among pigeons from three different sources. First, we used wild-caught Scottish and Irish ancestral-type Rock Doves (meaning those which putatively represent the undomesticated form of the species). These birds were sampled in various locations and then categorized geographically as belonging to either the "Outer Hebrides" (N = 44), "Inner Hebrides & Arran" (N = 34), "Highlands & Orkney" (N = 16), "Shetland" (N = 9) or "Cape Clear Island" (N = 3) populations ([Figure 1A](#) and [Table S1](#)). Next, we sampled birds in captive collections that were claimed to be ancestral-type Rock Doves (N = 48). These came from eight private and public zoological collections in the Netherlands. The provenance of all of these "Rock Doves" was unclear, and many of the collections are known to have indulged in recreational inter-species hybridization in the past (including with the Stock Dove *Columba oenas*), as is common in aviculture ([Ottenburghs et al., 2015](#)). Despite this, the captive "Rock Doves" conformed to the phenotype of the nominate subspecies *C. l. livia* ([Johnston and Janiga, 1995](#)). Finally, we used samples from pigeons of domestic origin, including fancy breeds (N = 38) as well as feral pigeons from the USA (N = 2), England (N = 26), Isle of Man (N = 3), and Lerwick, Shetland (N = 9) ([Table S1](#)). The captive/fancy domestic pigeon breeds have varying provenance, including various regions of continental Europe, Asia, Africa, and the Middle East. Wild Rock Doves were genetically differentiated from both the fancy/feral pigeons and captive "Rock Doves." This was supported by groupings resolved by Principal Component Analysis (PCA) based on genotype likelihoods at 2,583,745 polymorphic nuclear sites filtered for linkage disequilibrium (see [STAR Methods](#)), where feral pigeons from England, the Isle of Man, and Scotland were genetically closer to American feral pigeons, and fancy pigeons originating from various global regions, than their Scottish and Irish Rock Dove neighbors ([Figure 1B](#)). Nine feral pigeons from Lerwick in Shetland clustered with the fancy/feral pigeons rather than the Rock Doves with which they share their island, which provides further support to the idea of a wild/feral population split rather than, for example, geographic structure owing to different colonization waves or differentiation between feral populations. Furthermore, maximum likelihood estimation of individual ancestries using NGSadmix ([Skotte et al., 2013](#)) indicated that $K = 3$ was the optimal model fit according to the Evanno method ($K = 1$ to $K = 7$ tested), where the first group corresponded to wild Rock Doves, the second to fancy/feral pigeons, and the third to the captive "Rock Doves" ([Figures S1A and 3](#)). $K = 2$ had the second highest support and split the Rock Doves from all other birds

(Figure S1B). When depicted as a maximum likelihood phylogenetic tree using IQTree from a dataset of 871,968 SNPs (rather than genotype likelihoods) and 121 birds, both the Rock Dove clade and the fancy/feral/captive clade showed 100% bootstrap support (Figure S2). The identification of wild Rock Doves as a distinct genomic group, which is an outgroup to the clade containing all sampled domestic pigeons (including very ancient breeds such as the ice pigeon (Figure S1B)), provides evidence that remnant populations of ancestral-type Rock Doves persist (although more global sampling and the use of ancient DNA would be required to provide a more accurate understanding of ancestry, given the possibility of today's Rock Doves representing an ancient feral pigeon population). Feral pigeons grouping with domestic pigeons in both PCA, NGSadmix, and the phylogeny reveals a close relationship with the homing/racing pigeons that make up a significant portion of their ancestry (Johnston and Janiga, 1995). The captive "Rock Doves" provenance remains unclear. About a quarter of these birds were likely of domestic origin, grouping with the fancy/feral population according to PCA and NGSadmix. The remaining birds had an unknown ancestry component, potentially the result of extensive hybridization with different subspecies of Rock Dove and other pigeon species. Further genetic structure was evident within the wild Rock Doves, with a strong distinction between the Outer Hebridean birds and those from elsewhere in Scotland (Figure S3).

Genomic distinctiveness of Rock Doves from their feral domestic conspecifics was mirrored by morphological differences, though intermediates were observed (Figures 2A and 2B). We quantified morphological variation using PCA of 149 birds caught around the British Isles, including English feral pigeons ($n = 66$) and Rock Doves from Scotland and Cape Clear Island ($n = 83$). Morphological differences primarily relate to head shape—for example, Rock Doves have smaller ceres, and the angle of the head rising from the base of the bill is smaller than in feral pigeons (Tables S2 and S3, Figure 2A). Contrary to prior opinion, we show that these differences are sufficient to allow ancestral-type Rock Doves (at least in the British Isles) to be distinguished from feral pigeons in the field, even among those with similar plumage patterns (Figure 2C) (Blasco et al., 2014). These differences also support early suggestions (based on skeletal comparisons using very small sample sizes) that the Rock Dove was morphologically distinct from feral pigeons at the level of avian subspecies (Johnston et al., 1988). Individuals with intermediate morphology were more common in regions with more opportunities for hybridization (e.g., doves from the "Highlands & Orkney" region) (Hewson, 1967) (Figure 2B). The attenuation of a distinctive Rock Dove morphology in these regions suggests that hybridization is leading to a breakdown in the genetic integrity of the Rock Dove lineage, and will also complicate phenotypic identification of non-hybrids in such regions.

Despite being both genetically and morphologically differentiated from feral pigeons, the Scottish and Irish Rock Dove populations did show evidence of introgressive hybridization with feral pigeons, but to different extents. We characterized the pattern of gene flow each Rock Dove population has experienced from feral domestic pigeons. First, we used NGSadmix at $K = 3$ (the best supported value of K) to examine the level of mixed ancestry in each individual (using the LD-filtered dataset of genotype likelihoods at 2,583,745 polymorphic nuclear sites). This analysis revealed negligible admixture in the Outer Hebrides, some in the "Inner Hebrides & Arran" group and Shetland, and higher levels in all individuals sampled from Cape Clear Island and the "Highlands & Orkney" group (Figure 3). The NGSadmix plot also shows that, when assessing both the feral pigeons of Lerwick in Shetland and the "England & Isle of Man" group, Rock Dove ancestry was highest in the Isle of Man, Yorkshire, and Lerwick. At these sites, Rock Doves were present in the last century (Brown and Grice, 2005; Fr et al., 1949). The identification of introgression between feral pigeons and Scottish/Irish Rock Doves was further supported by TreeMix analysis (Pickrell and Pritchard, 2012), which identified gene flow between the domestic and Rock Dove clades. Examining multiple runs for each number of putative migration events suggested that four migration edges (three between the fancy/feral group and the Rock Dove group, and one within the Rock Dove group) best explained the sample covariance (Figure S4A). There was a notable lack of migration events to or from the Outer Hebrides group (Figure S4B). TreeMix also supported the topology of the maximum likelihood phylogeny.

Population-level, genome-wide estimates of introgression from feral pigeons into the Scottish/Irish Rock Doves were estimated using ABBA-BABA SNP patterns (using the genomic dataset of 871,968 SNPs, the sister species the Hill Dove *Columba rupestris* as an outgroup, and calculating Patterson's D -statistic using *Dtrios* in the *Dsuite* software package (see STAR Methods)). Introgression was detected from the "England & Isle of Man" feral pigeons into the "Inner Hebrides & Arran" ($D = 0.005$, $Z = 6.21$, f_4 -ratio = 3.86%, p -value < 0.001), Cape Clear Island ($D = 0.012$, $Z = 3.86$, f_4 -ratio = 10.35%, p -value < 0.001), "Highlands

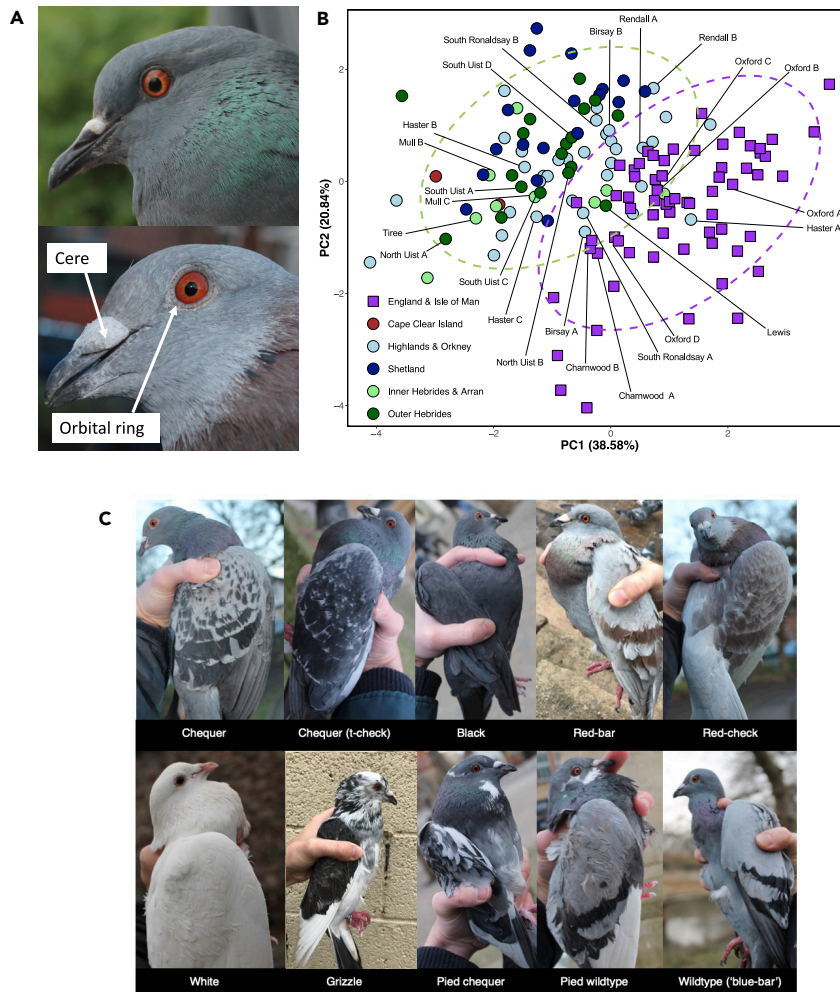


Figure 2. Rock Dove morphology

(A) The large cere, thick bill, flattened forehead and pale/thick orbital ring of the feral pigeon (bottom: from Leicester, England) distinguish it from the undomesticated Rock Dove (top: from the Inner Hebrides, Scotland).

(B) Rock Doves and feral pigeons in the British Isles are morphologically differentiated from each other, albeit with overlap, according to Principal Components Analysis of morphological variation based on six traits (wing length, cere width, tarsus length, bill length, head angle, and mass) in 149 birds. 95% confidence interval ellipses show feral pigeons (purple) and Rock Doves (green). Variance explained by PC1 and PC2 is shown in parentheses on the axes. Labeled individuals are birds that are also part of the genomic dataset.

(C) Although Rock Doves naturally exhibit only the “wildtype” (“blue-bar”) pattern, feral domestic pigeons can exhibit identical plumage, as well as a variety of other patterns and colors. All birds in the photographs are feral pigeons from Oxford and Leicester, England. Note the “feral pigeon” head shape is evident in all the birds pictured here, even the one with wildtype plumage.

& Orkney” ($D = 0.015$, $Z = 11.41$, f_4 -ratio = 12.59%, p -value < 0.001), and, to a lesser extent, Shetland ($D = 0.003$, $Z = 2.58$, f_4 -ratio = 2.11%, $p < 0.001$) populations relative to the Outer Hebridean Rock Doves. Whilst introgression detected in Rock Dove populations depends on the level of local contact with feral pigeons, the broader population genetic structure reflected the degree of geographic isolation from the core of the feral pigeon range in Great Britain. Genome-wide weighted F_{ST} estimates between the feral pigeons of “England & Isle of Man” and the Rock Doves of the British Isles were lowest for the more proximate regions (“Inner Hebrides & Arran” ($F_{ST} = 0.084$), “Highlands & Orkney” ($F_{ST} = 0.077$)) and highest for the most peripheral regions (Cape Clear Island ($F_{ST} = 0.148$), Shetland ($F_{ST} = 0.105$) and Outer Hebrides ($F_{ST} = 0.099$)).

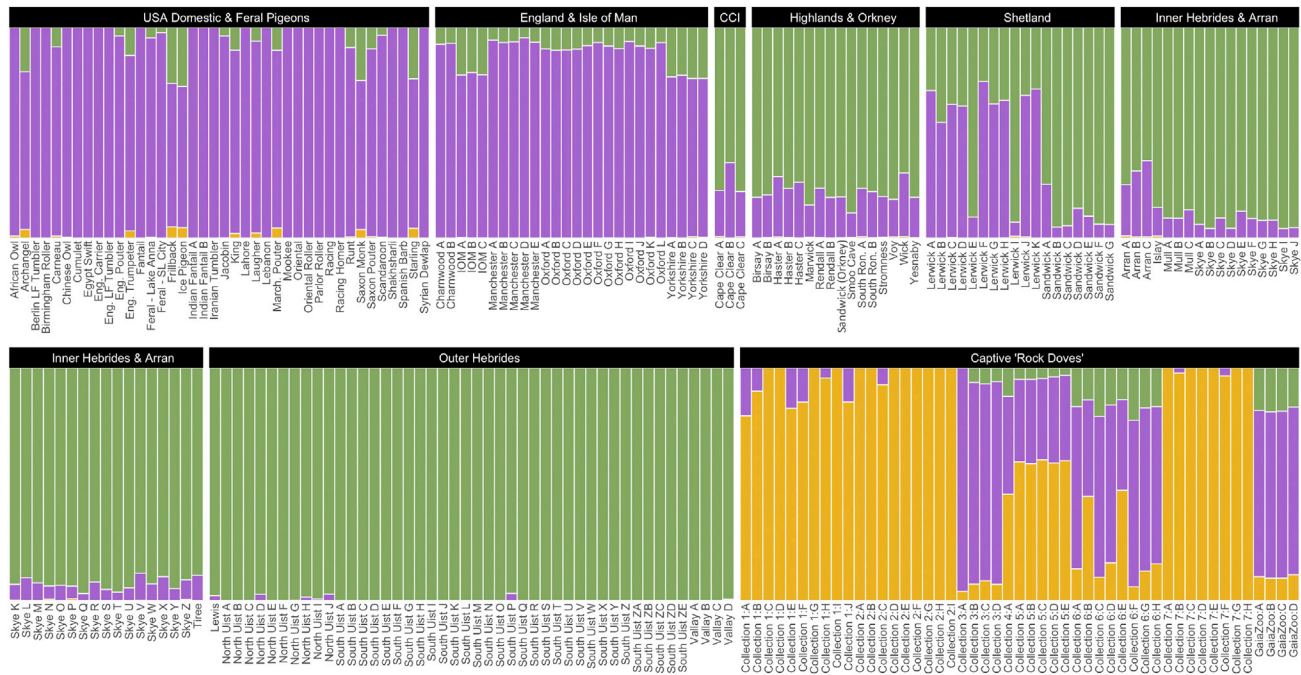


Figure 3. Maximum likelihood estimation of individual ancestries for feral/domestic pigeons, wild Rock Doves, and captive “Rock Doves”
 Calculated using NGSadmix using a linkage disequilibrium-filtered dataset of genotype likelihoods containing 2,583,744 nuclear sites, and the most likely value of $K = 3$ (see Figure S1A). There are variable levels of admixture between the feral/domestic pigeon group (purple) and different wild Rock Dove populations (green), with the least admixed Rock Dove individuals occurring in the Outer Hebrides. The third cluster (amber) was primarily found in captive “Rock Doves”, although some of these individuals also showed high levels of admixture from feral-domestic pigeons (purple).

The distribution of introgression across the genome was quantified using a D -statistic derivative, f_d , considered more robust to small sample sizes (here, windows of 40 informative SNPs) (see STAR Methods for more information regarding window-based analyses) (Martin et al., 2015). This revealed that, relative to the Outer Hebridean Rock Doves, feral pigeon introgression was distributed heterogeneously throughout the genome in the Rock Doves of the “Inner Hebrides & Arran,” “Highlands & Orkney,” Shetland, and Cape Clear Island (Figure 4A). The degree of genetic differentiation (F_{ST}) between each Rock Dove population and the feral pigeons was also distributed heterogeneously throughout the genome (Figure 4B). Windows of high introgression and differentiation were often tightly clustered; however, the position of most of these clusters varied between populations. The relative divergence of chromosome Z was higher than that observed in autosomes, a pattern seen in many avian divergence studies, and hypothesized to relate to the presence of a disproportionate number of genes involved in sexual selection and reproductive isolation, as well as the smaller effective population size of sex chromosomes and reduced recombination relative to the autosomes (Oyler-McCance et al., 2015; Sendell-Price et al., 2020). Most genes associated with peaks of introgression or differentiation differed between populations, reflecting the varied impacts of gene flow, drift, and selection (Table S4). Exceptions were the *CYBB* gene, associated with skeletal system development (Ng et al., 2015), which was associated with peaks of differentiation with the “Inner Hebrides & Arran,” “Highlands & Orkney” and Shetland groups; and *PAX9* and *NKX2-1*, which play roles in morphogenesis (Li et al., 2017; Puelles et al., 2000), associated with introgressed regions in Cape Clear Island and differentiated regions in Shetland. Given their known roles in shaping phenotypes, these and other shared genes could contribute to the morphological differentiation described here between Rock Doves and feral pigeons. Together, the results for individual levels of admixture, alongside population-level introgression statistics, support the identification of the Outer Hebridean Rock Doves as being the least introgressed. As the extent of hybridization across the global Rock Dove population appears to be very high (Baldaccini, 2020), the results described here mark the Outer Hebrides as a population of conservation significance. Populations from other parts of Scotland, particularly in the Highlands and Orkney, are more introgressed and unlikely to experience a reprieve from ongoing hybridization (Murton and Westwood, 1966). Without intervention, it is likely that Scottish Rock Doves will follow the fate of their now-extinct English and Welsh Rock Dove neighbors (Brown and Grice, 2005; Lovegrove et al., 1994). Overall, the presence of birds with

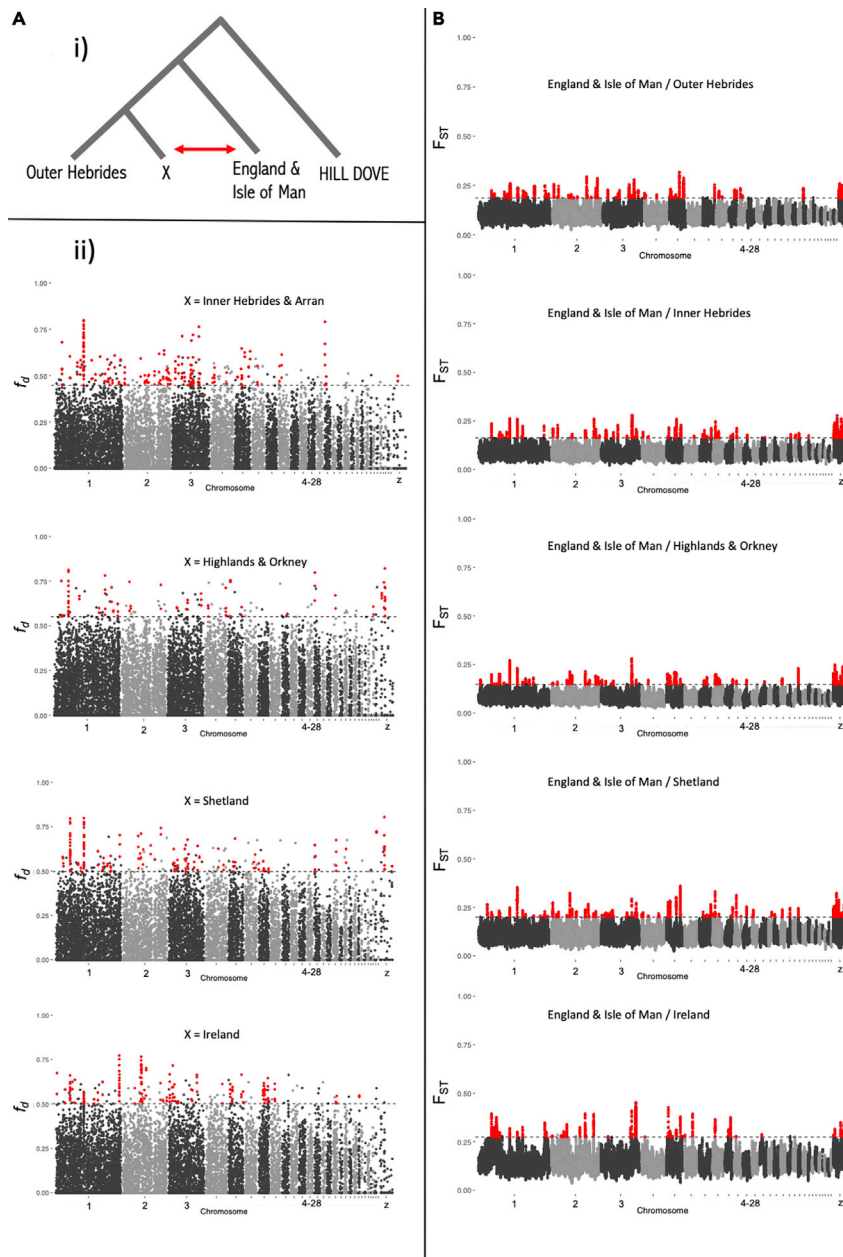


Figure 4. Genome-wide variation in introgression and population differentiation

(A) (i.) Tests for introgression (using f_d) included the Hill Dove *C. rupestris* as an outgroup, with fixed positions for the Outer Hebrides and “England & Isle of Man” populations. Estimates of introgression were made between the “England & Isle of Man” and each remaining Rock Dove population (X) (as indicated by the red arrow). f_d was calculated using a window size of 40 informative SNPs (i.e., SNPs that change the numerator of the statistic for a trio of populations) and step size of 20 SNPs, with the VCF of 871,968 SNPs. (ii.) The f_d statistic showed that introgression has occurred across the genome in all population combinations tested. Chromosomes are depicted in alternating colors. The dashed line indicates +3 standard deviations above the mean (points above this are classified as outlier windows) and red points indicate “outlier clusters” where outlier windows fall within 10 windows of one another.

(B) Pairwise F_{ST} comparisons of each Rock Dove population with the “England & Isle of Man” feral population showed that peaks of genetic differentiation occur throughout the genome. Chromosomes are depicted in alternating colors. The dashed line indicates +3 standard deviations above the mean (points above this are classified as outlier windows) and red points indicate “outlier clusters” (where outlier windows fall within 10 windows of one another). F_{ST} was calculated with a window size of 100,000 and a step size of 10,000 base pairs.

mixed ancestry, and the genome-wide nature of introgressed genetic material, highlights the ongoing genetic replacement of the ancestral-type lineage, and the real risk of their global extinction by hybridization.

DISCUSSION

The occurrence of extinction by hybridization is proliferating in an increasingly anthropogenic world, particularly involving feral domesticates and their wild relatives (Ottenburghs, 2021; Smith et al., 2022). Indeed, the wild-domestic interface is viewed as an increasingly important component of biodiversity (Clark et al., 2018). The Rock Dove populations we have identified have a primarily insular distribution, with populations on different islands have experienced varying levels of introgression ranging from the least admixed Outer Hebrides populations to sites where the ancestral-type Rock Dove may now be considered to be extinct as a distinct lineage (e.g., parts of Orkney). Most ancestors of domestic forms, such as junglefowl and wildcats, have primarily continental distributions and often occur as mixed wild-domestic populations (Senn et al., 2019; Wu et al., 2020). In contrast, Rock Doves present an unusual geographic arrangement of a wild-domestic interface with varying levels of isolation experienced, providing a tractable system to explore the spatial dynamics of incipient extinction by hybridization as a contributor to biotic homogenization and declining global biodiversity (Olden, 2006; Rhymer and Simberloff, 1996). The rarity of such populations is underscored by our discovery that putative captive populations of Rock Doves are genetically closer to domesticated populations.

The documentation of genetically and morphologically distinct Rock Dove populations provides an opportunity to understand the biology of an entity representing an undomesticated form, in its original setting (Smith et al., 2022). This will greatly enhance the utility of the domestic pigeon as a model organism. As well as its more traditional role in behavioral and neuroscience studies (Sasaki and Biro, 2017), the species is emerging as a valuable genetic model system (Shapiro et al., 2013) and, in its superabundant and cosmopolitan feral form, one used to understand microevolution in urban environments (Carlen and Munshi-South, 2021). In terms of the latter, understanding the wild context in which *Columba livia* evolved should facilitate the identification of traits that are derived adaptations for urban dwelling. More generally, a more nuanced understanding of how the species evolved, and the selection pressures it is exposed to in a contemporary environment, will enhance our ability to make inferences about the evolution of the behavioral, developmental, and genetic phenomena studied in domestic pigeons.

The Rock Dove offers an unprecedented opportunity to study an undomesticated form in a natural setting, providing the opportunity to enhance our understanding of the domestic pigeon as a model organism in the biological sciences. Concerningly, relict populations are at real risk of replacement with feral pigeons following extinction by hybridization. Future work should focus on comparing the ancestral-type, feral and captive domestic lineages of *C. livia* from around the world, to better explain phenomena studied in domestic pigeons, explore the extent to which Rock Doves represent the ancestral genetic variation, and unlock the potential of the species as a tractable model organism with which to explore domestication and hybridization.

Limitations of the study

To explore the ancestry of the sampled populations in more detail, and to place our study into a global context, more widespread geographic sampling would be required. The use of ancient DNA would also be useful. This could be used to establish the extent to which contemporary Rock Dove populations are representative of the original populations from which domestic pigeons originated, given the fact that domestication first happened thousands of years ago. Nevertheless, the Rock Doves identified here form an outgroup in phylogeny to domestic breeds originating from all over the world, including ancient breeds from the Middle East where the species was originally domesticated. They are, therefore, of significant conservation and scientific interest as wild representatives of the cosmopolitan domestic and feral pigeons. As well as this, it is probable that other populations of wild pigeons exist, particularly of subspecies other than the ancestral-type *livia*, which are relatively isolated from gene flow with contemporary feral and domestic pigeons. These could be in northern Africa for example (subspecies *gymnocyclus* or *dakhlae*). Again, more global sampling would be needed to establish this.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Data collection – Captive populations
 - Data collection – Wild populations
 - DNA extraction
 - Library preparation and sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Mapping
 - Genotype likelihood estimation and SNP calling
 - PCA analyses
 - NGSadmix
 - TreeMix
 - *D* statistics
 - F_{ST} calculation
 - Identifying genes associated with regions of interest
 - Phylogenetics
 - Phenotypic analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2022.104620>.

ACKNOWLEDGMENTS

Feral pigeon catching was carried out by WJS under Natural England license 2017-32742-SCI-SCI. Rock Doves were caught under British Trust for Ornithology (BTO) ringing licenses. Feather collection for DNA sampling was permitted under BTO Special Methods project 4211 with WJS (BTO Ringing Permit 6193) as project lead. We thank the citizen scientists, aviculturists, zoological collection staff, pest controllers, and landowners who either sampled birds themselves or enabled fieldwork to be carried out. I. Smith, J. Jones, A. Blackburn, C. Barker, G. Barker, S. Bayley, P. Harris, P. Ellis, G. Petrie, M. Gray, B. Ribbands, J. Bowler, A. Rhodes, B. Fudge, D. MacDonald, C. MacDonald, I. Thompson, Y. Benting, T. Millichap, S. Caygill, E. Prins, S. Bosch, F. Ketelslegers, G. Andeweg, P. Reinders, J. Bosch, J. Schrijvers, R. Van de Kerkhof, R. Marshall and G. Thompson all made especially valuable contributions to sample collection. We acknowledge the use of the Advanced Research Computing (ARC) facility at the University of Oxford, and Novogene and BGI for sequencing. We thank C. Quilodrán, T. Barraclough, R. Smith, I. Smith, B. Macted, D. Wawman, and S. Knowles for helpful comments on this article, and M. Shapiro, N. E. Baldacinni, A. Rhodes and D. Leech for encouraging words at the beginning of the project. E.C. Anderson, R. Pracana, and G. Van der Auwera provided valuable advice on data processing. WJS is funded by a DPhil studentship with the Edward Grey Institute of Field Ornithology. The study was further funded by a Heredity Fieldwork Grant from the Genetics Society (awarded to WJS), a Des Rubens and Bill Wallace Grant from the John Muir Trust (awarded to WJS), and a John Fell Grant from the University of Oregon (awarded to WJS, LAT, SK, ALF, and SMC). Finally, we thank Irby Lovette and three anonymous reviewers, whose comments allowed us to improve the article.

AUTHOR CONTRIBUTIONS

Conceptualization: WJS, LAT, SK, and SMC, Data curation: WJS, ASP, and CVDK, Formal analysis: WJS, ASP, SK, and MTJ, Funding acquisition: WJS, LAT, SMC, ALF, and SK, Investigation: WJS, ASP, and MTJ, Methodology: TMS and KCR, Project administration: WJS and CVDK, Resources: SMC, KCR, and BCS, Supervision: SMC, BCS, ALF, LAT, and SK, Visualization: WJS, MTJ, SMC, and ASP, Writing – original draft: WJS, SMC, and ASP, Writing – review & editing: WJS, MTJ, SMC, ASP, BCS, ALF, LAT, SK, KCR, TMS, and CVDK.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 31, 2021

Revised: May 27, 2022

Accepted: June 10, 2022

Published: July 1, 2022

REFERENCES

- Andrews, S. (2015). *FastQC*.
- Baker, K. (2016). Identification of European Non-passerines (British Trust for Ornithology).
- Baldaccini, N.E. (2020). The synanthropic status of wild rock doves (*Columba livia*) and their contribution to feral pigeon populations. *Riv. Ital. Ornitol.* 90, 479. <https://doi.org/10.4081/rio.2020.479>.
- Blasco, R., Finlayson, C., Rosell, J., Marco, A.S., Finlayson, S., Finlayson, G., Negro, J.J., Pacheco, F.G., and Vidal, J.R. (2014). The earliest pigeon fanciers. *Sci. Rep.* 4, 5971. <https://doi.org/10.1038/srep05971>.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Brown, A.F., and Grice, P.V. (2005). *Birds in England (T & A D Poyser)*.
- Carlen, E., and Munshi-South, J. (2021). Widespread genetic connectivity of feral pigeons across the Northeastern megacity. *Evol. Appl.* 14, 150–162. <https://doi.org/10.1111/eva.12972>.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., and Cresko, W.A. (2013). Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22, 3124–3140. <https://doi.org/10.1111/mec.12354>.
- Clark, N.J., Seddon, J.M., Šlapeta, J., and Wells, K. (2018). Parasite spread at the domestic animal-wildlife interface: anthropogenic habitat use, phylogeny and body mass drive risk of cat and dog flea (*Ctenocephalides* spp.) infestation in wild mammals. *Parasit Vectors* 11, 8. <https://doi.org/10.1186/s13071-017-2564-z>.
- Damas, J., O'Connor, R., Farré, M., Lenis, V.P.E., Martell, H.J., Mandawala, A., Fowler, K., Joseph, S., Swain, M.T., Griffin, D.K., and Larkin, D.M. (2017). Upgrading short-read animal genome assemblies to chromosome level using comparative genomics and a universal probe set. *Genome Res.* 27, 875–884. <https://doi.org/10.1101/gr.213660.116>.
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>.
- Durand, E.Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* 28, 2239–2252. <https://doi.org/10.1093/molbev/msr048>.
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.
- Fitak, R.R. (2021). OptM: estimating the optimal number of migration edges on population trees using Treemix. *Biol. Methods Protoc.* 6, bpab017. <https://doi.org/10.1093/biomethods/bpab017>.
- Fox, E.A., Wright, A.E., Fumagalli, M., and Vieira, F.G. (2019). ngsLD: evaluating linkage disequilibrium using genotype likelihoods. *Bioinformatics* 35, 3855–3856. <https://doi.org/10.1093/bioinformatics/btz200>.
- Fr, N., Botni, P., and Williamson, K. (1949). Polymorphism and breeding of the rock dove in the Faroe Islands. *Ibis* 91, 17–23. <https://doi.org/10.1111/j.1474-919X.1949.tb02233.x>.
- Francis, R.M. (2017). pophelper: an R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* 17, 27–32. <https://doi.org/10.1111/1755-0998.12509>.
- Goodwin, D., and Gillmor, R. (1970). *Pigeons and Doves of the World (Trustees of the British Museum). Natural History*.
- Hewson, R. (1967). The rock dove in Scotland in 1965. *Scott. Birds* 4, 359–371.
- Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. <https://doi.org/10.1093/molbev/msx281>.
- Huang, J., Liang, X., Xuan, Y., Geng, C., Li, Y., Lu, H., Qu, S., Mei, X., Chen, H., Yu, T., et al. (2017). A reference human genome dataset of the BGISEQ-500 sequencer. *GigaScience* 6, 1–9. <https://doi.org/10.1093/gigascience/gix024>.
- Huson, D.H., Richter, D.C., Rausch, C., DeZulian, T., Franz, M., and Rupp, R. (2007). Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinf.* 8, 460. <https://doi.org/10.1186/1471-2105-8-460>.
- Hutchinson, C. (2010). *Birds in Ireland (A&C Black)*.
- Johnston, R.F., Siegel-Causey, D., and Johnson, S.G. (1988). European populations of the rock dove *Columba livia* and genotypic extinction. *Am. Midl. Nat.* 120, 1–10. <https://doi.org/10.2307/2425881>.
- Johnston, R.F. (1992). Geographic size variation in rock pigeons, *Columba livia*. *Ital. J. Zool.* 59, 111–116. <https://doi.org/10.1080/11250009209386656>.
- Johnston, R.F., and Janiga, M. (1995). *Feral Pigeons (Oxford University Press)*.
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermini, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. <https://doi.org/10.1038/nmeth.4285>.
- Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., and Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15, 1179–1191. <https://doi.org/10.1111/1755-0998.12387>.
- Korneliusson, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: analysis of next generation sequencing data. *BMC Bioinf.* 15, 356. <https://doi.org/10.1186/s12859-014-0356-4>.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>.
- Li, C., Lan, Y., Krumlauf, R., and Jiang, R. (2017). Modulating Wnt signaling rescues palate morphogenesis in Pax9 mutant mice. *J. Dent. Res.* 96, 1273–1281. <https://doi.org/10.1177/0022034517719865>.
- Li, H. (2011). Improving SNP discovery by base alignment quality. *Bioinformatics* 27, 1157–1158. <https://doi.org/10.1093/bioinformatics/btr076>.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Lovegrove, R., Williams, G., and Williams, I. (1994). *Birds in Wales (Poyser)*.
- Malinsky, M., Matschiner, M., and Svardal, H. (2021). Dsuite - fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* 21, 584–595. <https://doi.org/10.1111/1755-0998.13265>.
- Martin, S.H., Davey, J.W., and Jiggins, C.D. (2015). Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol. Biol. Evol.* 32, 244–257. <https://doi.org/10.1093/molbev/msu269>.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Resour.*

20, 1297–1303. <https://doi.org/10.1101/gr.107524.110>.

Meisner, J., and Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics* 210, 719–731. <https://doi.org/10.1534/genetics.118.301336>.

Michaelsen, T., and Refvik, S. (2003). Feral rock doves breeding on the coast of Norway. *Ornis Nor.* 26, 119–124.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. <https://doi.org/10.1093/molbev/msaa015>.

Murton, R., and Westwood, N. (1966). The foods of the rock dove and feral pigeon. *Hous. Theor. Soc.* 13, 130–146. <https://doi.org/10.1080/00063656609476116>.

NCBI Genome Remapping Service. (<https://www.ncbi.nlm.nih.gov/genome/tools/remap>).

Ng, C.S., Chen, C.-K., Fan, W.-L., Wu, P., Wu, S.-M., Chen, J.-J., Lai, Y.-T., Mao, C.-T., Lu, M.-Y.J., Chen, D.-R., et al. (2015). Transcriptomic analyses of regenerating adult feathers in chicken. *BMC Genom.* 16, 756. <https://doi.org/10.1186/s12864-015-1966-6>.

Olden, J.D. (2006). Biotic homogenization: a new research agenda for conservation biogeography. *J. Biogeogr.* 33, 2027–2039. <https://doi.org/10.1111/j.1365-2699.2006.01572.x>.

Ottenburghs, J. (2021). The genic view of hybridization in the Anthropocene. *Evol. Appl.* 14, 2342–2360. <https://doi.org/10.1111/eva.13223>.

Ottenburghs, J., Ydenberg, R.C., Van Hooft, P., Van Wieren, S.E., and Prins, H.H.T. (2015). The Avian Hybrids Project: gathering the scientific literature on avian hybridization. *Ibis* 157, 892–894. <https://doi.org/10.1111/ibi.12285>.

Oyler-McCance, S.J., Cornman, R.S., Jones, K.L., and Fike, J.A. (2015). Z chromosome divergence, polymorphism and relative effective population size in a genus of lekking birds. *Heredity* 115, 452–459. <https://doi.org/10.1038/hdy.2015.46>.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. *Genetics* 192, 1065–1093. <https://doi.org/10.1534/genetics.112.145037>.

Picard. (<http://broadinstitute.github.io/picard/>).

Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* 8, e1002967. <https://doi.org/10.1371/journal.pgen.1002967>.

Puelles, L., Kuwana, E., Puelles, E., Bulfone, A., Shimamura, K., Keleher, J., Smiga, S., and Rubenstein, J.L.R. (2000). Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the genes *Dlx-2*, *Emx-1*, *Nkx-2.1*, *Pax-6*, and *Tbr-1*. *J. Comp. Neurol.* 424, 409–438. [https://doi.org/10.1002/1096-9861\(20000828\)424:3](https://doi.org/10.1002/1096-9861(20000828)424:3).

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. <https://doi.org/10.1086/519795>.

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*.

Rhymer, J.M., and Simberloff, D. (1996). Extinction by hybridization and introgression. *Annu. Rev. Ecol. Systemat.* 27, 83–109. <https://doi.org/10.1146/annurev.ecolsys.27.1.83>.

Sasaki, T., and Biro, D. (2017). Cumulative culture can emerge from collective intelligence in animal groups. *Nat. Commun.* 8, 15049. <https://doi.org/10.1038/ncomms15049>.

Schweizer, T.M., DeSaix, M.G., and Ruegg, K.C. (2021). LI-seq: a cost-effective, low input DNA method for whole genome library preparation. Preprint at bioRxiv. <https://doi.org/10.1101/2021.07.06.451326>.

Sendell-Price, A.T., Ruegg, K.C., Anderson, E.C., Quilodrán, C.S., Van Doren, B.M., Underwood, V.L., Coulson, T., and Clegg, S.M. (2020). The genomic landscape of divergence across the speciation continuum in island-colonising silvereyes (*Zosterops lateralis*). *G3* 10, 3147–3163. <https://doi.org/10.1534/g3.120.401352>.

Senn, H.V., Ghazali, M., Kaden, J., Barclay, D., Harrower, B., Campbell, R.D., Macdonald, D.W., and Kitchener, A.C. (2019). Distinguishing the victim from the threat: SNP-based methods reveal the extent of introgressive hybridization between wildcats and domestic cats in Scotland and inform future in situ and ex situ management options for species restoration. *Evol. Appl.* 12, 399–414. <https://doi.org/10.1111/eva.12720>.

Shapiro, M.D., Kronenberg, Z., Li, C., Domyan, E.T., Pan, H., Campbell, M., Tan, H., Huff, C.D., Hu, H., Vickrey, A.I., et al. (2013). Genomic diversity and evolution of the head crest in the rock pigeon. *Science* 339, 1063–1067. <https://doi.org/10.1126/science.1230422>.

Skotte, L., Korneliusson, T.S., and Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics* 195, 693–702. <https://doi.org/10.1534/genetics.113.154138>.

Smith, W.J., Quilodrán, C.S., Jezierski, M.T., Sendell-Price, A.T., and Clegg, S.M. (2022). The wild ancestors of domestic animals as a neglected and threatened component of biodiversity. *Conserv. Biol.* 36, e13867. <https://doi.org/10.1111/cobi.13867>.

Soraggi, S., Wiuf, C., and Albrechtsen, A. (2018). Powerful inference with the D-statistic on low-coverage whole-genome data. *G3* 8, 551–566. <https://doi.org/10.1534/g3.117.300192>.

Stringham, S.A., Mulroy, E.E., Xing, J., Record, D., Guernsey, M.W., Aldenhoven, J.T., Osborne, E.J., and Shapiro, M.D. (2012). Divergence, convergence, and the ancestry of feral populations in the domestic rock pigeon. *Curr. Biol.* 22, 302–308. <https://doi.org/10.1016/j.cub.2011.12.045>.

Thom, V.M. (2010). *Birds in Scotland* (Bloomsbury Publishing).

Wu, M.Y., Low, G.W., Forcina, G., van Grouw, H., Lee, B.P.Y.H., Oh, R.R.Y., and Rheindt, F.E. (2020). Historic and modern genomes unveil a domestic introgression gradient in a wild red junglefowl population. *Evol. Appl.* 13, 2300–2315. <https://doi.org/10.1111/eva.13023>.

Zheng, Y., and Janke, A. (2018). Gene flow analysis method, the D-statistic, is robust in a wide parameter space. *BMC Bioinf.* 19, 10. <https://doi.org/10.1186/s12859-017-2002-4>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Feathers from wild-living and captive <i>Columba livia</i>	See Table S1	N/A
Pectoral muscle tissue from freshly dead <i>Columba livia</i>	See Table S1	N/A
Chemicals, peptides, and recombinant proteins		
Glycogen	Life Technologies	Cat# R0561
Proteinase K	Strattech Scientific	Cat# K1037
Dithiothreitol	Merck	Cat# D9779-5G
Ethanol	Enzo Life Sciences	Cat# SV-39556-01
Sodium acetate	Merck	Cat# 32319-500G-R
Phenol/Chloroform/Isoamyl alcohol	Fisher Scientific	Cat# 327111000
Molecular Biology Grade Water	VWR International	Cat# 436912C
NaCl	Insight Biotechnology	Cat# SC-295832
Sodium dodecyl sulfate	Merck	Cat# L5750-500G
EDTA	Merck	Cat# 03620-50G
Tris-HCl	Promega UK	Cat# A2641
AMPure XP	Beckman Coulter	Cat# A63881
Critical commercial assays		
DNeasy Blood & Tissue Kit	QIAGEN	Cat# 69506
Tagment DNA TDE1 Enzyme and Buffer Kit	Illumina	Cat# 20034198
Qubit 2.0 Fluorometer	ThermoFisher	Cat# Q32850
TapeStation	Agilent	Cat# G2991BA
Deposited data		
Raw sequencing reads	This study	ENA: PRJEB52260
Software and algorithms		
FastQC v0.11.9	(Andrews, 2015)	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Trimmomatic v0.39	(Bolger et al., 2014)	http://www.usadellab.org/cms/?page=trimmomatic
ANGSD v0.933	(Korneliussen et al., 2014)	http://www.popgen.dk/angsd/index.php/ANGSD
VCFTools v0.1.16	(Danecek et al., 2011)	https://vcftools.github.io/index.html
Samtools v1.12	(Li et al., 2009)	http://www.htslib.org/
Bowtie 2 v2.4.1	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
GATK v4.1.8.1	(McKenna et al., 2010)	https://gatk.broadinstitute.org/hc/en-us
Picard v2.26.10	(Picard, http://broadinstitute.github.io/picard/)	https://broadinstitute.github.io/picard/Picard
PLINK v2.00	(Purcell et al., 2007)	https://zzz.bwh.harvard.edu/plink/
ngsLD v1.1.1	(Fox et al., 2019)	https://github.com/fgvieira/ngsLD

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PCAngsd v0.98	(Meisner and Albrechtsen, 2018)	http://www.popgen.dk/software/index.php/PCAngsd
NGSadmix v32	(Skotte et al., 2013)	http://www.popgen.dk/software/index.php/NgsAdmix
OptM v0.1.6	(Fitak, 2021)	https://cran.r-project.org/web/packages/OptM/index.html
CLUMPAK	(Kopelman et al., 2015)	http://clumpak.tau.ac.il/
DSuite v0.4	(Malinsky et al., 2021)	https://github.com/millanek/Dsuite
IQTREE v2.1.3	(Minh et al., 2020)	http://www.iqtree.org/
Stacks v 2.41	(Catchen et al., 2013)	https://catchenlab.life.illinois.edu/stacks/
Treemix v1.13	(Pickrell and Pritchard, 2012)	https://bitbucket.org/nygcresearch/treemix/wiki/Home
Dendroscope v3	(Huson et al., 2007)	https://uni-tuebingen.de/en/fakultaeten/mathematisch-naturwissenschaftliche-fakultaet/fachbereiche/informatik/lehrstuehle/algorithms-in-bioinformatics/software/dendroscope/
R v4.0.2	(R Development Core Team, 2010)	https://www.r-project.org/
Pophelper v2.3.1	(Francis, 2017)	http://www.royfrancis.com/pophelper/articles/index.html

Other

<i>Columba livia</i> reference genome colLiv2	(Damas et al., 2017)	GenBank: GCA_001887795.1
<i>Columba livia</i> reference genome Cliv_1.0	(Shapiro et al., 2013)	GenBank: GCA_000337935.1
<i>Columba livia</i> reads from fancy domestic pigeon breeds and USA feral pigeons	(Shapiro et al., 2013)	GenBank: PRJNA167554
NCBI Genome Remapping Service	(NCBI Genome Remapping Service, https://www.ncbi.nlm.nih.gov/genome/tools/remap)	https://www.ncbi.nlm.nih.gov/genome/tools/remap

RESOURCE AVAILABILITY

Lead contact

Requests for further information should be directed to and will be fulfilled by the Lead Contact, William Smith (william.smith@queens.ox.ac.uk).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The accession number for the sequencing data reported in this paper is publicly available (ENA: PRJEB52260). No original code was written for the analysis of this project. Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Our dataset contains 232 *Columba livia* individuals. Of these, 40 come from publicly available sequencing reads (NCBI: SRA054391) from fancy and feral pigeons (we also used one individual of the Hill Dove *Columba rupestris* for use as an outgroup in phylogenetic and introgression analyses) sampled in the USA as part of a previous study (Shapiro et al., 2013). A further 49 samples were from the feathers of birds claimed to be ancestral-type Rock Doves in both public zoological and private avicultural collections. The remaining 143 birds were sampled from feral pigeon and Rock Dove populations across the United

Kingdom, Isle of Man and Ireland (117 within the traditionally accepted range of the undomesticated Rock Dove according to existing literature, and 26 outside of it (Brown and Grice, 2005; Hutchinson, 2010; Lovegrove et al., 1994; Thom, 2010)). The Scottish and Irish samples were categorized according to geographical region for analyses. These categories were: 'Shetland', 'Highlands & Orkney', 'Inner Hebrides & Arran', 'Outer Hebrides' and 'Cape Clear Island'. Nine of eighteen birds caught in Shetland were feral pigeons and were excluded from introgression analyses to prevent a biased overestimate of the extent of gene flow into Shetland Rock Doves. Table S1 provides further detail on sample collection. All samples were stored in a freezer at -20°C before DNA extraction. All collection of samples responsible for the newly generated sequencing data was carried out between 2017 and 2020. Sampling and bird handling was approved by both the British Trust for Ornithology's Special Methods Technical Panel (SMTP) and the University of Oxford's Department of Zoology Animal Welfare and Ethical Review Body (AWERB). Sex of birds was unknown (the species cannot be reliably sexed in the field) and none were squabs (i.e. all were post-fledging age and no chicks in the nest were sampled).

METHOD DETAILS

Data collection – Captive populations

CVDK liaised with aviculturists responsible for seven different private avicultural collections, and WJS liaised with staff at GaiaZOO in the Netherlands. Ten flank feathers were taken from each of the 49 individual birds claimed to be ancestral-type Rock Doves within these collections. Birds were confirmed to conform to the phenotype expected of an ancestral-type Rock Dove (*C. l. livia*) according to existing literature (Goodwin and Gillmor, 1970).

Data collection – Wild populations

Birds were caught across various sites in the British Isles, by both WJS and other qualified British Trust for Ornithology bird ringers. Three flank feathers were taken from each individual bird. WJS liaised with pest controllers in Yorkshire, Manchester and the Isle of Man, and acquired twelve fresh feral pigeon carcasses (no birds were killed for this study). A further eight freshly dead birds, found opportunistically (e.g. casualties of window collisions), were sent to WJS by volunteers local to Rock Dove colonies.

Birds caught in 2019 were measured for biometric analysis. Birds were first classified as juveniles or adults based on plumage (Baker, 2016), and only adults ($n = 149$) were included in the analyzed dataset to avoid including juveniles and chicks at variable growth stages. Six morphological traits were recorded by WJS: Maximum wing chord and tarsus length to the claw were measured with a butted metal ruler ($\pm 1\text{mm}$); cere width and bill length (to feather edge) were measured with dial calipers ($\pm 0.1\text{mm}$); mass was measured with a digital pesola scale ($\pm 0.1\text{g}$) and head angle (from the top line of the bill to the frontal rise of the forehead) was measured with a computerized protractor ($\pm 1^{\circ}$) (Protractor v20.6.1) from a lateral-view photograph of each bird. Overlap between birds sampled for morphological and genetic data collection is highlighted in Figure 2.

DNA extraction

For the single carcass sample from Tiree, DNA was extracted from 25mg of pectoral muscle tissue using a Qiagen DNeasy Blood & Tissue Kit (Qiagen, Venlo, Netherlands), following the manufacturer's instructions. For feather samples, and the pectoral muscle samples from additional carcasses, DNA was extracted using the following phenol-chloroform extraction protocol. Samples were assigned randomly to an extraction batch, rather than extracting batches according to fieldwork location or putative Rock Dove/feral pigeon status. For each bird sampled, three to five mm from the end of each feather calamus was cut and split lengthwise with a scalpel on a clean laboratory tile to expose the tissue inside, then added to a microcentrifuge tube containing 250 μL of DIGSOL extraction buffer (0.02 M EDTA, 0.05 M Tris-HCl (pH 8.0), 0.4 M NaCl, 0.5% sodium dodecyl sulphate (SDS)), 4 μL of Proteinase K and 30 μL of 1 M Dithiothreitol (DTT) solution. Samples were incubated at 55°C for at least 5 h to overnight. Following incubation, 400 μL of phenol:chloroform:isoamyl alcohol (25:24:1) was added and samples centrifuged at 15,000 rpm for 15 min. The aqueous layer was transferred to a new microcentrifuge tube containing 35 μL of 3 M sodium acetate, vortexed for 10 s and chilled for 10 min at -20°C . DNA was precipitated with addition of 700 μL of ice-cold 100% ethanol and 2 μL of glycogen, chilled overnight at -20°C , centrifuged for 30 min at 15,000 rpm and the supernatant was removed. The precipitate was rinsed with 900 μL of ice-cold 70% ethanol, placed in a refrigerated centrifuge at 4°C for 30 min at 15,000rpm, the supernatant removed and the precipitated

DNA left to dry at room temperature. Once dried, the DNA was resuspended in 40 μL of Low-EDTA TE (Tris-EDTA) buffer (0.01 M Tris-HCL (pH 8.0), 0.0001 M EDTA).

Library preparation and sequencing

Library preparation and sequencing of the Tiree sample was conducted at BGI Genomics (Hong Kong). Library preparation used the MGI SP-100 automated system, and sequencing was conducted using the BGISEQ-500 platform (Huang et al., 2017), using paired end 150-bp sequence reads. Library preparation of all other samples was carried out by TMS using a protocol designed for increased efficiency when using low-input DNA (Schweizer et al., 2021). DNA samples were quantified using a Qubit 2.0 (Invitrogen, Waltham, Massachusetts, USA). Nine feather DNA samples with a concentration below 1 ng/ μL were concentrated from 40 μL total volume to 5 μL total volume using a 1:1 bead ratio with Ampure XP beads (Beckman Coulter, Brea, CA, USA) and quantified again using the Qubit. DNA samples were normalized to 2.5 ng/ μL , with an accepted range from 0.3 ng/ μL to 4.0 ng/ μL . 1 μL DNA was then fragmented and adapters were annealed (called Tagmentation) using Illumina's Tagment DNA TDE1 Enzyme and Buffer Kit. DNA then went through a two-step PCR amplification totaling 12 cycles to add unique Nextera dual indices (Illumina, San Diego, CA, USA) and amplify tagmented DNA, called an 'Indexing PCR' and 'Booster PCR'. Individual DNA libraries were then size selected and cleaned using a 0.7:1 bead to DNA ratio of Ampure XP beads, and eluted into 30 μL of 10 mM Tris-HCl. Individual cleaned DNA libraries were quantified using Qubit and pooled in equal amounts based on the library with the lowest total DNA yield. Ampure XP beads were used to complete a double size selection on pooled library to remove fragments outside of desired range of 320-500 bp, and library was eluted into 50 μL of 10mM Tris-HCl. Final library quantification was 1.55 ng/ μL (4.95 nM) and final TapeStation showed an average fragment size of 474 bp (insert size: 361bp). The library was sequenced on four Illumina HiSeq 4000 lanes (Illumina, San Diego, CA, USA) at Novogene USA, using paired end 150-bp sequence reads.

QUANTIFICATION AND STATISTICAL ANALYSIS

Mapping

The quality of all sequencing reads was assessed using FASTQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Based on quality reports we removed the first 10bp of each read using Trimmomatic v0.39 (HEADCROP:10) (Bolger et al., 2014). Trimmed reads were aligned to the *Columba livia* chromosomal genome assembly 'colliv2' (NCBI Assembly GCA_001887795.1) using Bowtie 2 v2.4.1 (Langmead and Salzberg, 2012) with end-to-end alignment and default settings (allowing for a maximum of two mismatches in the seed (-n 2)) and resulting BAM files sorted using SAMtools v1.12 (samtools sort) (Li et al., 2009).

Genotype likelihood estimation and SNP calling

We carried out SNP calling on all 233 birds (i.e. including all *Columba livia* individuals plus *Columba rupestris*, the Hill Dove) using HaplotypeCaller in GATK v4.1.8.1 with the *Columba livia* genome assembly v2 (NCBI Assembly GCA_001887795.1), the genotyping_mode 'DISCOVERY' and a confidence threshold of 20 (-stand-call-conf 20). This was to generate data for IQTREE, TreeMix and Dsuite, which take a VCF as an input. SNP-based analyses can be significantly affected by variable sequencing depth in different individuals. To achieve a sensible range of sequencing depths of 4-10x, we used DownSampleSam in Picard (java -jar picard.jar DownsampleSam) to reduce the sequencing depth of the Tiree sample ($p = 0.2$), and of all the USA Domestic & Feral Pigeon samples ($p = 0.5$). The outputted VCF file was filtered with VCFtools v0.1.16 (Danecek et al., 2011) to get only biallelic sites (-min-alleles 2, -max-alleles 2) remove indels (-remove-indels) and only include sites where the minor allele count was $\leq 5\%$ (-maf 0.05); genotype quality was >20 (-minGQ 20), and where genotypes were called in at least 50% of individuals (-max-missing 0.5). Individual data missingness was then calculated using VCFtools v0.1.13 (Danecek et al., 2011) and individuals with $>50\%$ missing data removed (-remove) from the VCF file. The resulting VCF contained 121 individuals and 871,968 biallelic Single Nucleotide Polymorphisms (SNPs). For TreeMix analysis, we filtered this file using PLINK (Purcell et al., 2007) with the settings -indep-pairwise 1000kb 1 0.8, to reduce the impact of linkage disequilibrium. This led to the retention of 675,581 SNPs in the linkage-disequilibrium-filtered VCF.

For the population structure analyses, we used ANGSD v0.933 (Korneliussen et al., 2014) to estimate genotype likelihoods (-doGlf) using the chromosomal *Columba livia* genome assembly v2 (colLiv2) (NCBI Assembly GCA_001887795.1) and the GATK genotype likelihood model (-GL 2), inferring major and minor

from GL data (`-doMajorMinor 1`) and including only proper pairs (`-only_proper_pairs 1`), using `qscore` recalibration with the BAQ model (`baq -1`) (Li, 2011). We set thresholds of base quality of at least 30 (`-minQ 30`) and mapping quality of at least 25 (`-minMapQ 25`), and discarded 'bad' reads (`-remove_bads 1`) and those that did not map uniquely (`-uniqueOnly 1`). We repeated this genotype likelihood estimation twice; once for all 232 *Columba livia* individuals (using `-minInd 116`), and once for all 108 of the British and Irish undomesticated Rock Doves (117 birds from Rock Dove regions minus nine individuals from Lerwick in Shetland which were feral pigeons) using `-minInd 54`. Because the files were destined for NGSadmix and PCAngsd, which required low linkage between sites, we then ran `ngsLD` (Fox et al., 2019) on both output files to estimate pairwise linkage disequilibrium. We filtered the outputted file using the `prune_graph.pl` script provided for `ngsLD`, with the settings `-max-kb-dist 5` and `-min_weight 0.2`. For the file of 232 birds, the outputted list contained 7,742,273 sites. For the file of 108 birds, it contained 7,353,590 sites. We then repeated the generation of genotype likelihoods as before, but restricted to the low-LD sites using the `ANGSD -sites` flag. Following this filtering for linkage disequilibrium, the final 232 bird genotype likelihood file contained 2,583,745 sites, and the 108 bird British/Irish genotype likelihood file contained 2,554,226 sites.

PCA analyses

We examined the patterns of population structure by Principal Component Analysis (PCA) of genotype likelihoods using PCAngsd's (Meisner and Albrechtsen, 2018) 'pcangsd.py' script for the full dataset of all 232 *Columba livia* individuals. We repeated this separately for the genotype likelihood file containing only the 108 birds from the British Isles (Figure S3A).

NGSadmix

We performed estimation of individual admixture proportions for all 232 birds using the genotype likelihood file and NGSadmix (Skotte et al., 2013). We tested K values of between 1 and 7 to determine which grouping had the most support, using sites where at least 116 individuals had NGS data (`-minInd 116`). For each value of K we conducted 10 runs, and summarized these runs using CLUMPAK (Kopelman et al., 2015) with the Evanno method (Evanno et al., 2005). $K = 3$ had the highest ΔK value (meaning $K = 3$ is the uppermost level of structure) between 1 and 7. We plotted the $K = 3$ and $K = 2$ (which had the second highest ΔK value) results using `pophelper` (Francis, 2017) in R (R Development Core Team, 2010). We repeated this for the file with 108 of the British Isles birds using `-minInd 54` (Figure S3). Here, $K = 2$ had the highest Delta K value (Figure S3B).

TreeMix

We used TreeMix (Pickrell and Pritchard, 2012) as a further test of hybridization. TreeMix models population relationships as a tree in which positions can be connected by migration edges representing genetic admixture between lineages. We used the LD-filtered VCF of 675,581 SNPs and converted it to a 'treemix' file using the `Stacks -treemix` flag (Catchen et al., 2013). We then performed the analysis varying the number of migration edges (M) from one to twelve. To identify the number of migration events that optimized the fit of the data to the tree, we used `OptM` (Fitak, 2021). `OptM` allows one to judge the optimal number of migration edges by outputting the Δm statistic, which represents the second order rate of change across values of M , peaking at the correct number of simulated migration edges. We ran TreeMix 50 times ($i = 50$) per M value, using *Columba rupestris* as a root and a randomly selected window size (k) of between 100 and 1000 SNPs (varying in 50 SNP increments). Varying the window size is necessary to prevent the standard deviation across runs being zero, which would make Δm undefined. `OptM` identified four migration events as optimizing the fit of the data to the tree (Figure S3C).

D statistics

Using `Dsuite` (Malinsky et al., 2021) and the VCF of 871,968 SNPs generated by GATK, along with the `-g` flag instructing the program to use the genotype likelihoods in the VCF, we carried out D statistic introgression analyses for the geographical groupings of Scottish and Irish Rock Doves. We carried out both genome-wide and within-genome introgression estimates. We first assessed the genome-wide levels of feral introgression using the D statistic (Durand et al., 2011; Patterson et al., 2012; Soraggi et al., 2018). To do so, we used the `Dtrios` program of `Dsuite` with all British and Irish populations (excluding the Lerwick feral pigeons from the Shetland population to prevent a biased overestimate of introgression between feral pigeons and the Shetland Rock Doves). `Dtrios` outputs a D statistic, Z-score, p-value and an f_4 -ratio (of admixture) for

each trio of populations (setting populations as 'P1', 'P2' and 'P3'). If a D statistic is not zero and has a Z score greater than 3, then it is deemed significant (Zheng and Janke, 2018). A positive D statistic implies introgression between P2 and P3, and a negative D statistic implies introgression between P1 and P3.

Following this genome-wide 'global' introgression analysis, to determine how the pattern of introgression varied across the genome, we calculated windowed D statistics using *Dinvestigate* (also from *Dsuite*). Rather than the D statistic itself, we used f_d , which is also used to detect loci that have experienced introgression, but is more reliable when basing calculations on smaller genomic regions, as is the case with window-based analyses (Martin et al., 2015). Unlike with *Dtrios*, for *Dinvestigate*, in each calculation P1, P2, P3 (phylogeny positions) and the outgroup must be set manually (see Figure 4A). We kept P1 as the 'Outer Hebrides' group (shown to have experienced negligible introgression by NGSadmix and TreeMix), P3 as the "England & Isle of Man" feral pigeon group and the outgroup as *Columba rupestris*. We then varied P2 (shown as 'X' in Figure 4A) and carried out four calculations, enabling us to judge introgression between feral pigeons and all of the Scottish and Irish Rock Dove populations where significant gene flow with feral pigeons was identified by *Dtrios* (Highlands & Orkney, Shetland, Inner Hebrides & Arran and Cape Clear Island), relative to the Outer Hebrides. In each case, the analysis quantified the pattern of feral pigeon introgression with the population set as P2. We used a window size of 40 informative SNPs (i.e. SNPs that change the numerator of the statistic for a trio of populations) and a step size of 20 (-w 40,20), and plotted f_d along the genome. We plotted the midpoints of genomic windows in R, and classified 'outlier regions' as those where the f_d value was more than 3 standard deviations above the mean. We classified 'outlier clusters' as being present where outlier regions were within 10 windows of each other.

F_{ST} calculation

We estimated both genome-wide and window-based genetic differentiation between the 'England & Isle of Man' feral pigeons and each of the Scottish/Irish Rock Dove populations (excluding Lerwick's feral pigeons from the Shetland population to prevent a biased underestimate of differentiation between them and the feral pigeons). Firstly, we generated a list of variable sites. To do so, we ran ANGSD using a list of all the 'England & Isle of Man' birds plus all the birds in the comparison population (meaning that F_{ST} calculation was carried out five times to compare the feral pigeons to Rock Doves from the Outer Hebrides, Inner Hebrides & Arran, Highlands & Orkney, Shetland and Cape Clear Island) with the following settings: -remove_bads 1, -only_proper_pairs 0, -trim -0, -minMapQ 30, -minQ 20, -minInd 10, -baq 1, -doMaf 1, -doMajorMinor 1, -doPost 2, -GL 2 and -minMal 0.01. We then extracted the first two columns from the maf.gz output file and indexed this file of variable sites using ANGSD with the 'sites index' flags. Using ANGSD, we next calculated allele frequency likelihood for both the 'England & Isle of Man' population and the comparison population using the reference assembly instead of an ancestral state reference, and the flags -remove_bads, -only_proper_pairs, -trim 0, -minMapQ 30, -minQ 20, -baq 1, -dosaf 1 and -gl 2. The outputted saf.idx files were then used to calculate the folded site frequency spectrum (using -fold 1) with realSFS, outputting folded.sfs files. Following this, we were able to use realSFS fst index with the saf.idx files for both populations and -sfs with the folded.sfs file to output F_{ST} on a per site basis with -fstout (giving a fst.idx file). Finally, we could get global F_{ST} estimates with realSFS fst stats, and use the fst.idx file to get a global.fst file. We also calculated Fst estimates in a sliding window using realSFS fst stats2 and the fst.idx file with a window size of 100,000 and a step size of 10,000 (which led to a broadly similar window size to *Dsuite's* 'informative SNP's method). We plotted the midpoints of each window along the genome and classified 'outlier regions' as being 3 standard deviations above the mean. We classified 'outlier clusters' as being regions where the outliers fall within 10 windows of each other.

Identifying genes associated with regions of interest

We mapped the genome annotation of the scaffold-based Cliv_1.0 assembly (Shapiro et al., 2013) to the chromosomal (unannotated) assembly colliv2 using NCBI's Genome Remapping Service (<https://www.ncbi.nlm.nih.gov/genome/tools/remap>). The subsequent gff file contained the annotations associated with Cliv_1.0 mapped to the colliv2 genome. We first filtered this gff for the mRNA feature (which contains the names of genes and their location along the genome). We then added two columns to the list of 'outlier clusters' of f_d (i.e. the *Dinvestigate* output). These columns were called scan_start and scan_end, and correspond to 100,000 less than and 100,000 greater than the start and end of the window respectively. In R, we scanned for genes in the gff that fell within our f_d scanning windows. We repeated this for F_{ST} outlier clusters.

Phylogenetics

We converted the VCF of 871,968 SNPs into a file in the PHYLIP format using Stacks `-phylip` (Catchen et al., 2013). This file was used to infer a maximum likelihood phylogenetic tree using variable sites with IQ-TREE 2, rooted using *Columba rupestris* (Minh et al., 2020). The best fitting model for this data matrix was selected automatically by IQ-TREE 2, enforcing (-m) ascertainment bias correction (+ASC) to correct for branch lengths in the absence of constant sites (Kalyaanamoorthy et al., 2017) and using the Bayesian information criterion, and was determined to be TVM + F + ASC + R3 (where TVM is the transversion model with unequal base frequency, F means that empirical base frequencies were used and R3 means that a FreeRate model of rate heterogeneity across sites with three categories was employed). A 1000 non-parametric bootstrapped re-sampling approach was taken to infer support values at internal nodes in the tree (Hoang et al., 2018). The tree was visualized using Dendroscope (Huson et al., 2007). The most basal node shows the split between the Rock Doves and the group containing both contemporary feral and domestic pigeons and captive 'Rock Doves' from private and public avicultural collections. Each of these clades has robust bootstrap support (Figure S2).

Phenotypic analysis

The six univariate morphological traits were standardized and log-transformed to conform to assumptions of normality and homogeneity of variance, and population differences between birds caught in putative Rock Dove regions and feral pigeons compared with t-tests (Table S3). A Principal Component Analysis (PCA) including all six morphometric traits and 149 adult birds was conducted with the function `prcomp` in the base R package (R Development Core Team, 2010). The first two Principal Components (PC's) collectively explained 59.42% of the variance (PC1 38.58% and PC2 20.84%). PC1 described increasing cere width, tarsus length, wing length and head angle, with feral pigeons having larger PC1 scores than Rock Doves. PC2 largely described bill length, with higher PC2 values corresponding to larger bills (Table S2).