

On the Origin and Spread of Feral Pigeons

George Pacheco^{1✉}, Filipe G. Vieira¹, Michael D. Martin¹, Morten Tange Olsen¹, Pavel Hulva¹, Tânia de Freitas Raso¹, Peter Njoroge¹, Concepción Salaberria¹, Isabel López-Rull¹, Carles Lalueza-Fox¹, Oscar Ramírez¹, María C. Ávila-Arcos¹, Patricia Rosas Escobar¹, Rui Faria¹, Miguel Carneiro¹, Graciela Sotelo¹, Jóhannis Danielsen¹, Nizar Haddad¹, Fares Khoury¹, Roi Dor¹, Ali Halajian¹, María Belén Arias¹, Oliver Krone¹, Susanne Auls¹, Sampath S. Seneviratne¹, Kajanka Mathiaparanam¹, Michael Bunce¹, Megan L. Coghlan¹, Jon Fjeldsø¹ & M. Thomas P. Gilbert^{1✉}

The rock pigeon (*Columba livia* Gmelin, 1789) is presumed native to the Mediterranean, Saharo-Arabian and Eastern Oriental regions, and is believed to have been domesticated in the Middle East in the early Neolithic. At some point during the domestication process, the first feral pigeons arose, whose populations subsequently undertook a remarkable expansion that has resulted in them being found today across almost the entire global urban landscapes. Indeed the spread of these feral birds has been so prolific, that it raises questions about whether any true wild rock pigeon colonies still exist, or whether they have been admixed with, or even fully replaced, by feral birds? While several studies have investigated the complex evolutionary history of pigeon breeds, none have yet addressed the question of pigeon feralisation, and how this evolutionary process might be jeopardizing the species' status as a wild entity. In this study, we generated and analysed a genomic dataset produced using the Genotyping-by-Sequencing (GBS) method of 450 feral pigeons sampled across 41 worldwide localities. Our analyses reveal that the global feral pigeon population can be divided into four major groups, each exhibiting different levels of genetic diversity and contamination with domesticated genotypes. We also find signs of strong population structure, including very divergent clades of what seems to be relatively wild populations. Lastly, we find evidence of human-mediated dispersal through past colonial links.

¹Section for Evolutionary Genomics, The GLOBE Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. ¹Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5–7, 1350 Copenhagen, Denmark. ¹NTNU University Museum, Norwegian University of Science and Technology, Trondheim, Norway ¹Department of Zoology, Charles University, Prague, Czech Republic. ¹Departamento de Patologia, Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, Brazil. ¹Ornithology Section, Department of Zoology, National Museums of Kenya, Nairobi, Kenya. ¹Centro de Investigación en Ecosistemas, Universidad Nacional Autónoma de México, Michoacán, México. ¹Departamento de Ecología Evolutiva, Museo Nacional de Ciencias Naturales, Madrid, Spain. ¹Avian Evolution Node, Department of Zoology and Environment Sciences, University of Colombo, Colombo, Sri Lanka. ¹Institute of Evolutionary Biology, Universitat Pompeu Fabra, Barcelona, Spain. ¹Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK. ¹Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Vairão, Portugal. ¹Institute of Evolutionary Biology, Department of Experimental and Health Sciences, University, Pompeu Fabra, Spain. ¹Departamento de Biología, Faculdade de Ciências, Universidade do Porto, Porto, Portugal. ¹University of the Faroe Islands, Tórshavn, Faroe Islands. ¹National Center for Agricultural Research and Extension, Al-Baqah, Jordan. ¹Department of Biology and Biotechnology, American University of Madaba, Madaba, Jordan. ¹Department of Zoology, Tel Aviv University, Tel Aviv, Israel. ¹Natural History Museum, Imperial College of London, London, United Kingdom. ¹Department of Biodiversity, Turfloop Campus, University of Limpopo, Polokwane, South Africa. ¹Department of Wildlife Diseases, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany. ¹Vetgenomics SL, Edifici Eureka, Campus UAB, Barcelona, Spain. ¹Trace and Environmental DNA (TyEnD) Laboratory, Department of Environment and Agriculture, Curtin University, Perth, Australia.

✉Correspondence should be addressed to ganpa@aqu.dtu.dk (G.P.) & tgilbert@sund.ku.dk (M.T.P.G.)

Archaeological evidence suggests that the rock pigeon (*Columba livia* Gmelin, 1789), and in particular the (*C. l. livia*) subspecies, was first domesticated during the Neolithic period in the Middle East, probably via a commensal pathway². While it was initially exploited as a source of food and fertiliser, later on, the extent of its service to humankind spanned a wider variety of roles, including incorporation into religious rituals, a tool for communication, a source of medicine, and even as a navigation aid³. Furthermore, in addition to its practical functional roles, and in parallel with many other domestic animals such as dogs, chickens and cats, the eighteenth century witnessed an explosion of interest in the development of so-called fancy breeds. Such interest led to the establishment of numerous pigeon breeds, of which over 230 are currently recognised by the *American National Pigeon Association* (NPA; www.npaua.com). Artificial selection in modern breeds resulted in a truly fabulous amount of phenotypic diversity, which has long attracted the attention of scholars, and even formed a cornerstone of Darwin's nascent thoughts on his famous theory concerning the evolutionary processes [1].

The history of the pigeon domestication has also been tightly coupled with a correlative evolutionary process—feralisation. The process of feralisation is thought to have begun through domestic pigeons escaping from captive stocks (kept within Europe, North Africa and Western Asia). As to the ecological niche of their wild ancestors, feral pigeons tend to utilise hardscape habitats as urban analogues of rocks/cliffs (Lundholm and Richardson, 2010) and become a synurbic species (Francis and Chadwick, 2012). As with other feralised domesticates, the species experienced an extreme ecological range expansion much later in time, when during the modern colonial period they were transmitted to, and subsequent release across, almost all continents of the world, successfully populating extensive urban areas. Therefore, at present, the feral pigeon is ubiquitous across the world's urban landscapes, where it is often considered a pest species requiring active management. Furthermore, given that both domestic and feral pigeons can almost certainly still interbreed with their wild ancestor, it has been proposed that in regions of co-occurrence, the wild rock pigeon gene pools (and thus their integrity as a natural species) might be to some extent contaminated with domestic genotypes as has been demonstrated for other domestic species^{5,6}.

Although studies aiming to shed light on the genomic relationships among pigeon breeds have been conducted^{7–9}, fundamental questions concerning today's feral pigeon populations have yet to be addressed in depth. These include i) which breeds principally contributed to the formation of the feral pigeon populations, and ii) which are the genomic relationships among these populations. Additionally, in light of the global expansion of both domestic stocks and feral populations during the past few centuries, and a third key question is iii) whether the wild rock pigeon has undergone a process of genomic extinction as seen with the wild ancestors of other domesticates^{6,10}.

In this study, we employed the Genotyping-by-Sequencing (GBS) method to generate a genomic dataset for 450 free-living pigeons from 41 localities covering a worldwide distribution, as well as XXX pigeons representing other subspecies/species^{??}. We use this dataset to reconstruct the phylogenetic relationships among these populations, as well as to investigate their patterns of both genetic structure and diversity revealed by analyses of MDS, Admixture and population genetics statistics.

We hypothesize that populations inhabiting remote localities within the believed natural range will show the lowest levels of contamination with domestic genotypes, while those populations inhabiting urban localities within the believed natural range will have moderate signals of admixture with domesticates. Besides, those populations outside the natural range will show the highest levels of influence by domestic genotypes, and, given that these populations were formed through human-mediated dispersals, we also expect that there will be a signal linking former colonies-colonizers relationships (e.g. London and Johannesburg).

Results

Sampling effort. To investigate the genomic patterns of current pigeon populations of different evolutionary histories, we intentionally targeted our sampling to cover four distinct categories (Figure 1). Furthermore, to help root the evolutionary relationships between these groups, we also included a small number of individuals representing the *Columba livia* intermedia (Strickland, 1844) subspecies. Specifically in this regard, we sampled five populations from Sri Lanka, where two of them were from urban localities (Colombo and Trincomalee), one was from a Conservation National Park (Pigeon Island) and two others were captive populations maintained by local breeders (Wattala and Wellawatte) (Supplementary Fig. 1). To check for data reproducibility, we sequenced two of the samples twice (Tehran_16-GBS and Perth_02-GBS) to serve as replicates. Finally, to serve as external outgroups, we also generated data from five samples of (*Columba palumbus* Linnaeus, 1758) captured in Copenhagen (Denmark), one captive sample of (*Streptopelia risoria* Linnaeus, 1758), and additionally incorporated previously published whole genome resequence data from a *Columba rupestris*⁸. (we also included the WGS library to serve as another replicate) (Supplementary Spreadsheet).

Sequencing data and filtering. We generated 1,062,089,639 reads using the (GBS) method¹¹ from all 488 samples. The percentage of GBS chimeric reads removed ranged from 0.034% to 4.65% (mean of 0.69%) per sample. Based on the presence/absence heatmap, we excluded 10 samples alongside the 6 blanks from downstream analyses (Supplementary Fig. 2). The sequencing depths across the regions of interest ranged from 0.069X to 19.97X (mean of 2.32X) (Supplementary Spreadsheet), with 68.86% of the predicted cut-sites yielding no data for any sample. This pattern is not unexpected with GBS data and derives from differences in sequencing effectiveness at different sites in the genome, and has also been observed in other GBS data from pigeons⁹. Moreover, based on the Global Depth (GD) results (Supplementary Fig. 3), we elected to run subsequent downstream analyses on only sites that have GD of at least 150X.

Population genetics statistics. Hybridisation between domestic lineages and their respective wild ancestors is an issue of conservation importance as the genomic integrity of wild lineages can be disturbed by contamination with domestic genotypes, either through direct admixture with purebred lineages or through indirect contact with feral stocks^{5,12}. To shed light on the levels of potential domestic contamination, as well as isolation among the sampled populations, we calculated several population genetics estimates for the different pigeon populations as well as F_{st} levels between them. First, we calculated the observed levels of heterozygosity (H_o) for each sample present in Dataset 1 (all samples). Among the free-living pigeon individuals, the H_o ranged from 0.1157%

to 0.2747% (average of 0.2231%) (Supplementary Spreadsheet), with considerable variation within each population and the occurrence of several outliers (Supplementary Fig. 5). On the other hand, the H_o ranged from 0.1818% to 0.2389% (average of 0.2176%) among the five *C. palumbus* samples.

Moreover, we calculated the nucleotide diversity (π), Watterson's Θ (Θ_w), and Tajima's D across the pigeon genome for the 35 populations that had five or more individuals (the Crete population was also included due to its relevance), as these genetic estimates only apply for population data (for the 2 replicate samples only the ones with higher coverage were used) (Supplementary Fig. 6). Π ranged from 0.001764 to 0.003080 (average: 0.002584; *C. palumbus*: 0.002427); Θ_w ranged from 0.001507 to 0.003168 (average: 0.002483; *C. palumbus*: 0.002644) and the Tajima's D ranged from -0.643739 to 0.854540 (average: 0.248505; *C. palumbus*: -0.413195). Interestingly, the values calculated for the five *C. palumbus* individuals in our study were lower than expected for a wild species, but we believe that the fact we analysed synanthropic individuals might have prevented us from accessing the true values for this species (Supplementary Spreadsheet).

We next elected to calculate the pairwise F_{st} between populations to provide preliminary insights into the relationships between the different populations. A heatmap plotted using this data, after excluding the populations of Tel Aviv Colony and *C. palumbus* that showed extremely high levels of divergence (Supplementary Fig. 7), clearly shows that three clusters are recovered. The first contains populations from relatively isolated regions within the natural range (Pigeon Island, Trincomalee, Torshavn, Crete, Sardinia, Vernelle, Wadi Hidan, Abadeh and Tehran). The second contains populations from mostly Europe and former colonies (Tel Aviv, Barcelona, Guimaraes, Lisbon, Berlin, Prague, Denver, London, Perth, Salvador, Copenhagen, Johannesburg and Salt Lake City). The third contains populations from Mexico (Santiago, San Cristobal de las Casas, Monterey, Tlaxcala de Xicohtencatl and Mexico City).

Overall, we see a wide range of genetic diversity levels among the sampled populations, similar to both domesticated (e.g. duck breeds: 0.0020 to 0.0028) and wild species (e.g. mallard: 0.0040)13. Tajima's D values show a general pattern of positive values, indicating population contractions, except for the populations from Tehran, Abadeh, Wadi Hidan, and Sardinia. Taken together, these results seem to group the current pigeon diversity into four groups, depending on the levels of genetic diversity (high, intermediate, or low), while the group with high levels of genetic diversity can be further subdivided depending on Tajima's D values (positive or negative). These groupings are further supported by the F_{st} distances, which clearly form groups around the populations with both intermediate and high genetic diversity.

Even though the levels of genetic diversity can give initial valuable insights regarding the past history of a given population, it is also important to infer the evolutionary relationships among the current pigeon populations. For that, we conducted the first phylogenetic reconstruction for non-domestic pigeon populations. Specifically, we first performed a Neighbour-joining (NJ) phylogenetic analyses based on Dataset 1 to confirm the appropriateness of the outgroups available (Supplementary Fig. 8). Having confirmed that *Columba rupestris* is the most appropriate outgroup, we proceeded with a Maximum-likelihood (ML) phylogenetic analysis based on Dataset 2 (excluding *S. risoria* and *C. palumbus*). Overall,

the ML phylogeny reconstructed reveals a strong phylogenetic structure across the different pigeon populations studied (Fig. 1). Despite the low bootstrap values (BS) recovered for the internal nodes (probably because the evolution of pigeons is not a truly bifurcating process), we delineated six main groups on our phylogeny based on geographical distribution and biological assumptions; Group A: The Pigeon Island and Trincomalee *C. l. intermedia* populations; Group B: Abadeh, Tehran, Crete, Sardinia, Vernelle and the Faroe Islands; Group C: Tel Aviv, Tel Aviv Colony and Wadi Hidan populations; Group D: Nairobi, Colombo, Lahijan, Nowshahr, Wellawatte and Isfahan populations; Group E: Guimaraes, Barcelona, Lisbon, Salvador, Tatui, Denver, Santiago, Tlaxcala de Xicohtencatl, Mexico City, Monterey and San Cristobal de Las Casas populations; Group F: Jihlava, Prague, Berlin, Salt Lake City, Johannesburg, London, Cambridge, Perth and Copenhagen. Interestingly, the four individuals belonging to the Wattala population do not form a monophyletic group and are found scattered across the phylogeny.

It is noteworthy that Group A, as well as both the Faroe Islands and the Tel Aviv Colony populations, have 100% of BS, with the former also having a very elongated branch length. Moreover, all the three replicates (namely, Crupestris_01, Tehran_16 and Perth_02) behave as expected, all clustering with 100% of BS. Furthermore, although both the eight Wattala and Wellawatte captive populations were located within the Colombo region, these populations do not form a distinct clade together with the Colombo population with some Wattala individuals sitting on a rather different region of the phylogeny. As for the three captive populations included in our study, all 6 individuals belonging to the Tel Aviv Colony population form a clade with 100% of BS and phylogenetic affinity with the Tel Aviv and the Hadi Hidan populations. The branch of this cluster is the longest across the entire phylogeny, in agreement with the fact that this colony has been maintained inbreed for several generations.

Phylogenetic relationships among feral pigeon populations.

Population structure among pigeon populations.

Contribution of pigeon breeds to current non-domesticated populations.

Discussion

Methods

Sequencing data generation and processing.

Data analysis.

Population genetics statistics.

Phylogenetic reconstruction.

Inference of Population Structure.

Contribution of pigeon breeds to current non-domesticated populations.

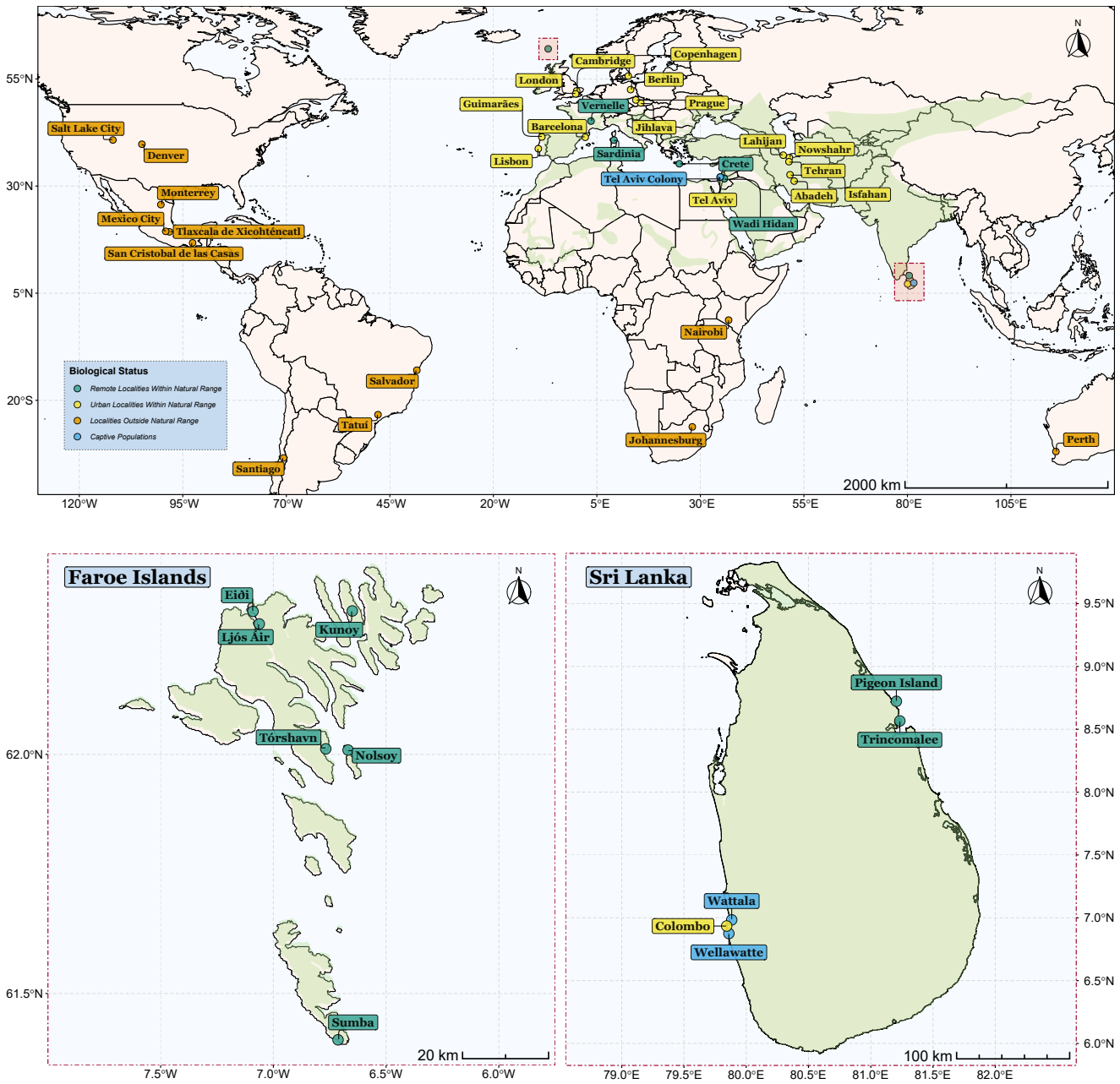
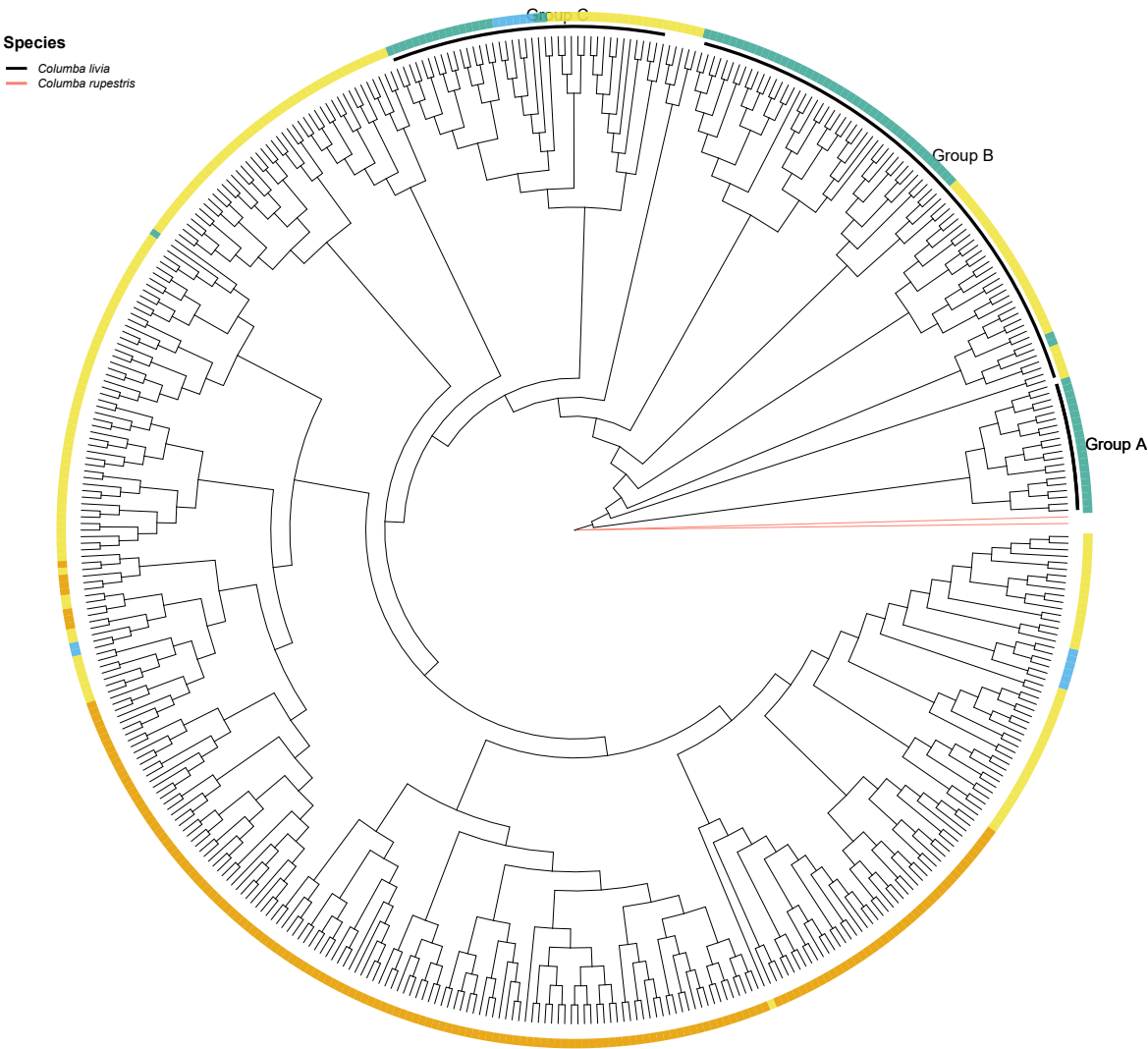


Fig. 1 Sampling effort. Map showing sampling localities. Samples were divided into four categories: a) remote localities within natural range (green); b) urban localities within natural range (yellow), localities outside natural range (orange); d) captives (blue). The supposed natural range of *Columbia livia* is depicted through a light green shade.



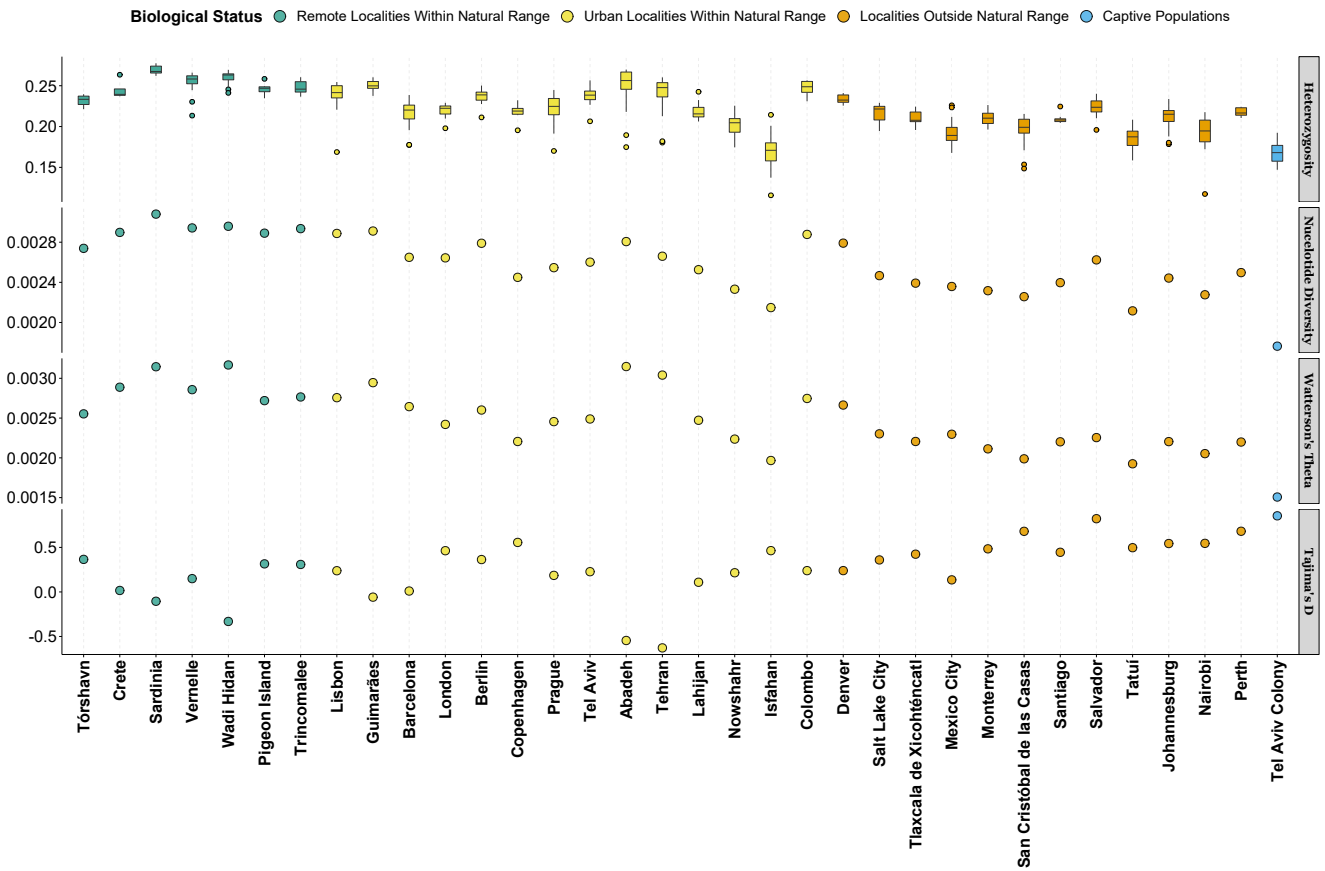


Fig. 3 Population genetics estimates. a) Proportion of observed heterozygous sites. b) Nucleotide Diversity (π). c) Watterson's Theta (θ_w). d) Tajima's D. All absolute values can be found in the Supplementary Spreadsheet.

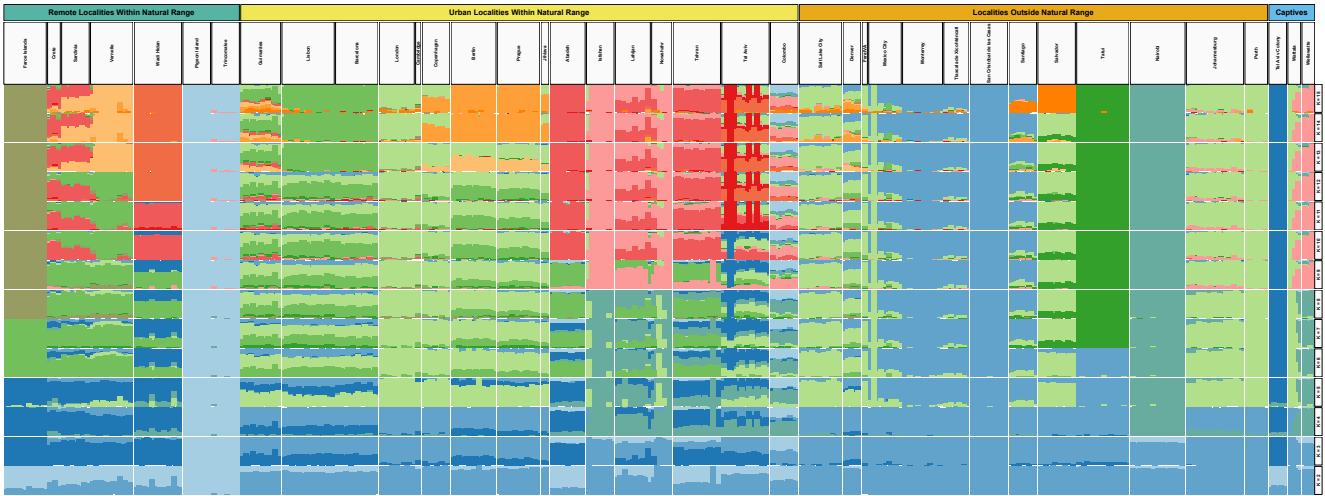


Fig. 4 Estimation of Admixture proportions. Individuals are represented by columns, while rows depict the Admixture proportions based on the assumption of different numbers of ancestral populations ($K = 2 - 15$). Individuals are sorted by sampling locality (light grey upper labels) and grouped per biological status.

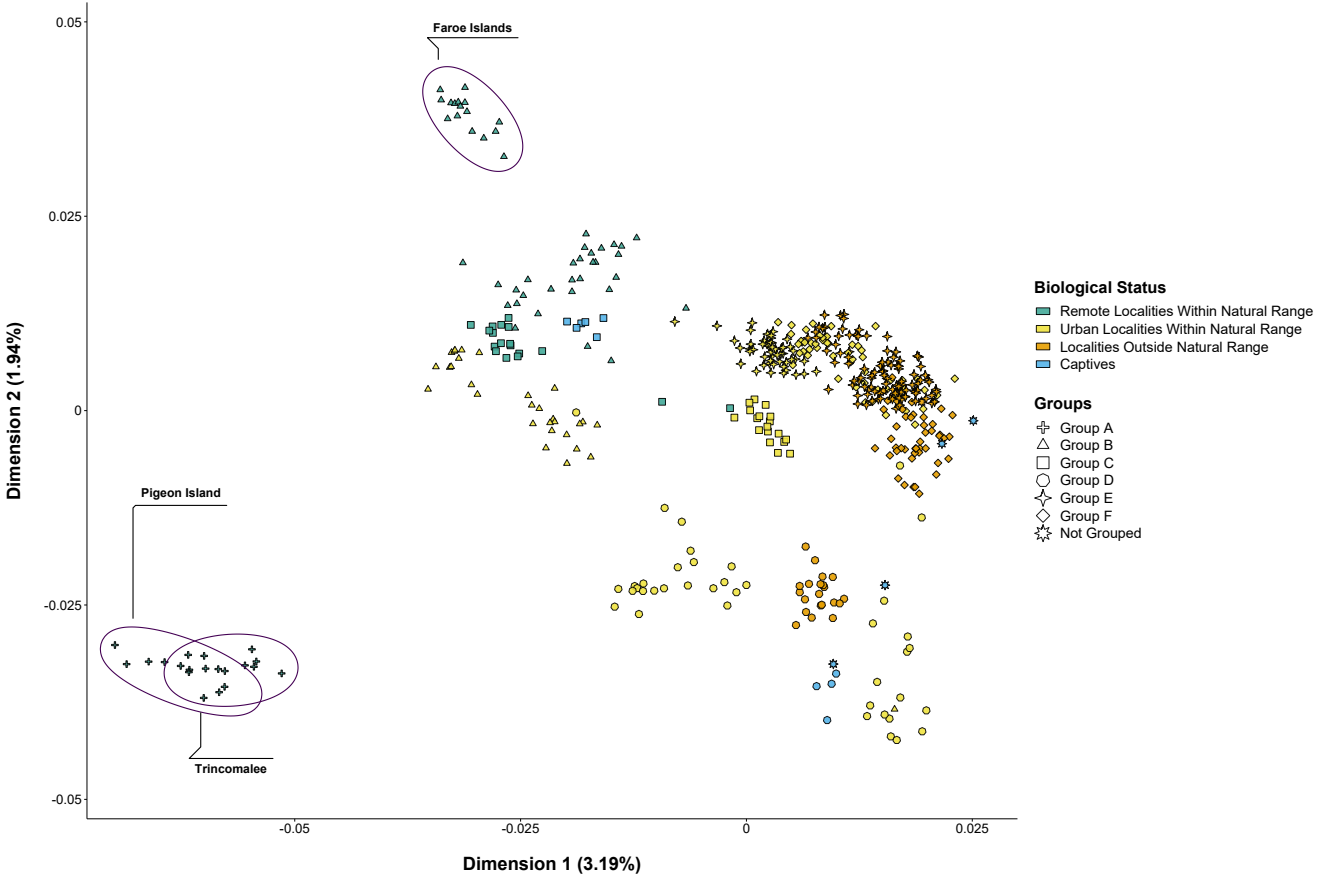


Fig. 5 Multidimensional scaling analysis. Dimensions 1 and 2 are plotted and each point on the plot represents a single individual. The ellipses represent the rough distribution of the two most distinct groups.

References

1. Pacheco, G., van Grouw, H., Shapiro, M. D., Gilbert, M. T. P. & Vieira, F. G. Darwin's Fancy Revised: An Updated Understanding of the Genomic Constitution of Pigeon Breeds. en. *Genome Biology and Evolution* **12** (ed Storz, J.) 136–150. ISSN: 1759-6653. <https://academic.oup.com/gbe/article/12/3/136/5735467> (2021) (Mar. 2020).

Acknowledgements

We would like to thank our local lab managers Charlotte Hansen, Pernille V. S. Olsen and Tina B. Brand at the Centre for GeoGenetics for their prompt support during the execution of the project. We are grateful to the Cornell University Biotechnology Resource Center for its genotyping services, especially to Sharon E. Mitchell and all lab technicians that worked on this project. Moreover, we deeply thank Gary Jakeman and Kristian Murphy Gregersen for their fieldwork assistance regarding the sampling in England and Denmark, respectively. We also thank Vladimir Orduña for his willingness to let us sample some of the Mexico City pigeons kept at his lab facilities.

Data Availability

All demultiplexed GBS sequencing data is publicly available at SRA (Project Number: PRJNA495951), as well as additional data uploaded to the University of Copenhagen's long term storage (https://sid.erda.dk/wsgi-bin/lis.py?share_id=aKqQoJvH4Y).

Code Availability

All codes required to reproduce the computational analyses are available from GitHub (<https://github.com/g-pacheco/FeralPigeonGenomics>).

Author Contributions

M.T.P.G. conceived the project and obtained financial support. M.T.P.G., G.P. and F.G.V. designed the study. G.P. led the project. M.T.P.G., G.P., M.T.O., T.dF.R., P.H., P.N., C.S., I.L.R., S.S.S., K.M., C. L.-F., G.S., R.F., J.D., J. F., N.H., F.K., R. D., A.H., M.B.A. M. C. A.-A. and P. R. E. contributed to sampling. M.D.S. collected and provided the breed samples. G.P. performed the vast majority of DNA extraction and QC. K.M. performed DNA extraction and QC on Sri Lanka samples. G.P. and F.G.V. conducted the computational analyses assisted by M.D.M. G.P., F.G.V., M.T.O. and M.T.P.G. interpreted the results. G.P. wrote the first draft of the manuscript with great input from M.T.P.G. and F.G.V. All authors critically reviewed and approved the final manuscript.

Funding

This project was funded by Lundbeck Foundation (award R52-5062) and European Research Council (Consolidator grant 681396) granted to MTPG. G.P. was supported by a Danish Government Scholarship and Tuition Fee Waiver Grant provided by the Danish Ministry of Science, Innovation and Higher Education and subsequently by a Ciência Sem Fronteiras Full PhD Abroad Scholarship (Grant Number: 201761/2014-9) provided by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) in Brazil.

Competing Interests

We have no competing interests.