Conservation genetics for the critically endangered Great Green Macaw

Salomé López Serna

Supervisors: Prof Jim Groombridge (main supervisor, Kent), Dr Dave Roberts (secondary supervisor, Kent), Dr Simon Tollington (external supervisor, Nottingham-Trent University)

Durrell Institute of Conservation and Ecology (DICE)

School of Anthropology and Conservation

University of Kent, UK

June 2023

Thesis submitted for the degree of Master's by Research

AUTHOR DECLARATION

I can confirm that this thesis and the work presented here are my own and represent the results of my original research. This work has not been submitted, in whole or in part, in any previous application for a degree, and was conducted while in candidature for a research degree at the University of Kent.

I conducted the laboratory work under the supervision of Jim Groombridge and the analysis and writing of the manuscript with his advice and that of Simon Tollington. Where I have consulted previous work published by others it has been duly attributed, and while recognizing the help and guidance of both my advisors and their support, this thesis is entirely my own work.

ACKNOWLEDGEMENTS

This research was supported by the British Ecological Society Small Grant Project SR20/1598, and by a Darwin Fellowship granted by the Darwin Initiative, Project: DARFW055 awarded to Jim Groombridge. I wish to thank both funders, without their support this project and my postgraduate studies would not have been possible. Thank you to the Macaw Recovery Network, Zooaves and the European Endangered Species Program for providing the samples for this project. Special thanks to Dr Sam Williams, Director of the Macaw Recovery Network, for his continued support of my interest in parrot conservation and research, and to Mario Jimenez, Macaw Recovery Network field team leader, for his commitment to obtaining wild samples and for teaching me field monitoring techniques for

the Great Green Macaw. I´m grateful for support from Anna Jemmet and Rebecca Louch of the Conservation Genetics lab at DICE for their guidance, experience, and constant willingness to help me with my analysis. Finally, many thanks to my lab mates at DICE, especially to Lucy and Pam, for their support and for making my lab work a great experience. The laboratory work and data analyses were carried out at the DICE Conservation Genetics Laboratory, University of Kent, with the addition of a short period of lab work at the University of Costa Rica (in the laboratory of Dr Eric Fuchs).

COVID DECLARATION

I started my research in January 2022, and by this year, most of the Covid restrictions had been lifted. I was able to develop my research for my thesis project normally and I did not find any part of my process hampered by the Covid 19 global pandemic. I was able to develop the objectives I had set out from the beginning of the project.

ABSTRACT

Ex-situ and captive breeding populations are important strategies for endangered species conservation, and the role of genetics is increasingly being recognized in their management. Furthermore, patterns of genetic variation and structure can provide managers with information to support conservation strategies. Molecular markers can be used to assess the status of captive populations, help prevent biodiversity loss and inbreeding and understand genetic diversity in wild and captive populations. We used newly developed microsatellite markers to examine genetic structure and diversity in 4 captive populations and one wild population of the critically endangered Great Green Macaw. We did not find significant differences in expected and observed heterozygosity and allelic richness between all the populations evaluated and found the value for expected heterozygosity in all of them within the range of other macaw and parrot species evaluated. We found genetic structure when we evaluated the five populations together which largely corresponded to three clusters formed by Costa Rican, European and Colombian samples, and found further fine genetic structure when we evaluated Costa Rican samples and the European samples independently. Additionally, we used the microsatellite marker set to determine relatedness between founders of one of the captive breeding populations and to evaluate the relatedness and genetic diversity in the Costa Rican captive and release populations. Our results contribute to the understanding of the genetic diversity of the species, and they can be used to galvanize the management of captive breeding and release populations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Genetic diversity is one of the three levels of biodiversity, in addition to species and ecosystem diversity (Verma, 2016). It is influenced by gene flow, population isolation, and genetic drift, and together they shape the genetic diversity and structure of wildlife populations (Allendorf, et al., 2013). The distribution of genetic diversity in a population is determined by large- and fine-scale spatial and temporal factors (Hofreiter & Stewart 2009; Manel, et al., 2003), and in species with wide distribution ranges, genetic diversity and structure may be associated with local adaptations shaped by diverging environments over time (Papadopulos et al., 2014) or lack of gene flow due to isolation by distance or physical barriers (Garnier et al., 2004, Herman et al., 2022, Frantz et al., 2010). However, either way, conserving genetic diversity at a species level is important to preserve a species' evolutionary potential (Eizaguirre & Baltazar-Soares, 2014). Furthermore, a growing body of scientific literature has established associations between genetic diversity and key factors related to a species' long-term persistence including reproductive success, viability/survival, disease/pathogen resistance and gamete quality (DeWoody et al., 2021), further highlighting the importance of conservation genetics as a management tool for conserving threatened species (Willi et al., 2022).

Captive populations are an important conservation resource to tackle species biodiversity loss because they serve as insurance populations for endangered species since they are protected from threats that cause wild populations to decline such as habitat loss, poaching, predation, and disease (IUCN 2014). Captive populations are also important source populations for reintroductions or conservation translocations to increase the size of wild populations (IUCN/SSC 2013; Seddon et al. 2014; Brichieri-Colombi et al., 2019; Frankham et al. 2010).

Captive populations usually start with a small number of founders and thus are vulnerable to some of the same genetic risks as small wild populations, namely loss of genetic diversity, accumulation of levels of inbreeding, and problems associated with inbreeding depression such as reduced fitness of individuals and increased susceptibility to infectious diseases (Aguiar et al., 2018; Alledrof et al., 2012; Farquharson et al., 2021; Frankham et al., 2010; Groombridge et al., 2012; Willoughby et al., 2015). Understanding the genetic composition of captive populations can provide information for management strategies to avoid these genetic risks. Furthermore, multiple captive populations of the same species can be managed using assisted gene flow informed by genetic data and can increase overall population viability by managing these *ex-situ* populations as a single metapopulation (Gooley et al., 2022). Genetic management has been successfully implemented in captive populations of birds (Alcaide, et al., 2010), mammals (Ramirez et al., 2006), fish (Fisch et al., 2015) and reptiles (Moore et al., 2008; Miller et al,. 2009) and integrating molecular genetic data into *ex situ* management is being increasingly recognized as an important tool over traditional studbook management (Attard et al., 2016; Hogg et al., 2017).

For successful captive breeding management, knowledge of pedigree and genealogical data is necessary to achieve the aim of retaining genetic diversity and limiting inbreeding (Ivy et al., 2009; Frankham, et al., 2017). To achieve this, populations have been traditionally managed using studbooks, but this approach has two major recognized pitfalls: assumed unrelatedness between founders and incompleteness in the known relationships between population members (Ballou, 1983). Unintentional accumulative errors in pedigree-managed populations limit the use of this strategy and can cause a decrease in fitness of the managed population (Hammerly, et al., 2013; Hammerly, et al., 2016). Integrating genetic data can

197 resolve both issues and accomplish more robust inferences and it has been done for some
198 managed species (McGreevy, et al., 2011; Henkel et al., 2012; Ferrie et al., 2013; Hammerly
199 et al., 2016; Overbeek et al., 2020), but research considering pedigree and genetic data are
200 still scarce (Ayala-Burbano, 2020).

201 Understanding the genetic composition of captive populations is also important in the context
202 of reintroduction of captive bred individuals into the wild. Success for reintroduced
203 populations will not only depend on habitat suitability, demographic and social factors for
204 the establishment and persistence of the population over time (Ewen et al., 2012), but also
205 on the genetic fitness and evolutionary potential of the established reintroduced population
206 (Pacioni, et al., 2013, Pacioni et al., 2020).

207 ICUN reintroduction guidelines indicate that genetic information about the founder
208 individuals is essential because two risks associated with translocation failure stem from it:
209 a risk of inbreeding depression caused by a reintroduction bottleneck, and low genetic
210 variation caused by uninformed selection of individuals for reintroduction, both of which
211 may hinder survival probability (and consequently, reintroduction success) and longer-term
212 adaptation to environmental change by a reintroduced population (IUCN/SSC, 2013).

213 Reintroduced populations often represent populations that have gone through multiple
214 population bottlenecks, which makes them inherently susceptible to inbreeding and loss of
215 genetic variation (Mock et al., 2004; Frankham 2005; Groombridge et al., 2012; White et al.,
216 2017). Since the genetic diversity and structure of the reintroduced population depends on
217 diversity within its source population, genetic results provide conservation managers with
218 baseline information to understand the genetic make-up of the potential founder population,
219 and present a starting point for future monitoring, assessment and genetic management aimed

to preserve genetic variability and avoid inbreeding depression in the wild (Frankham et al., 2002; De Barba et al., 2010; Schreier et al., 2015; Carroll et al., 2018).

The Great Green Macaw (*Ara ambiguus*) is the second-largest New World psittacine (Monge et al., 2010), and it is classified as critically endangered by the International Union for Conservation of Nature - IUCN Red List (BirdLife International, 2020). Since 2000, the population has been declining rapidly, being classified as Vulnerable in 2000, Endangered in 2005 and Critically Endangered in 2020. It is distributed in Central America, from Honduras to Panama, Colombia and Ecuador (BirdLife International, 2020) (Figure 1). Ecuador is the only country containing the subspecies *Ara ambiguus guayaquilensis*, while the other subspecies *Ara ambiguus ambiguus* has a distribution across Honduras, Nicaragua, Costa Rica, Panama and Colombia (Fjeldså et al., 1987). The global population size is estimated in the band between 500-1000 mature individuals, with the subpopulation of the Caribbean slope of southern Nicaragua/north-eastern Costa Rica initially estimated at ~160 mature individuals, but with a later estimation of around 485 individuals in Costa Rica (Lewis et al., 2022) . Habitat loss and poaching for the pet trade are the main causes of the species' population decline (BirdLife International, 2020). In Central and South America, deforestation due to agriculture, cattle, illegal plantations, road expansion, mining and logging are the most significant threats (BirdLife International, 2020). The Great Green Macaw has a strong dependency on the Almedro tree (*Dipteryx oleifera*), a widely distributed canopy tree species, for availability of nest sites and as a food source (Monge et al., 2003), however this tree species is heavily logged for the wood trade in Colombia and Costa Rica and is considered to be a key reason for the dramatic decline of the Great Green Macaw. To date, several captive breeding programs have been set up for the species, including the

Macaw Recovery Network (MRN) and Zooaves *ex situ* populations in Costa Rica, the Parque de la Conservacion *ex situ* population in Colombia and the European Endangered Species Program (EEP), in concordance with one of the IUCN conservation actions proposed for the species (Birdlife International, 2020).

By estimating genetic diversity and structure in captive and wild populations of the Great Green Macaw, we can better understand and conserve this species, and move towards firstly, establishing a genetically informed captive breeding program that better safeguards the species and maintains this critical aspect of its diversity, and secondly, informing reintroductions efforts that contribute to a cohesive conservation program for the species focused on increasing population size.

In this study we aimed to (i) optimize a set of microsatellite loci developed *de novo* for the Great Green Macaw, (ii) determine levels of genetic diversity and genetic structure within and between four captive and a single wild population of the species, and (iii) apply the genetic data to help guide future management of captive populations and to inform future reintroduction planning for this critically endangered species.

METHODS

Samples

We sampled 149 Great Green Macaw individuals from several captive and wild populations including the captive-breeding populations managed by the Macaw Recovery Network and by Zooaves in Costa Rica, the captive population at Parque de la Conservacion in Colombia, the European Endangered Species Program managed by the European Association of Zoos and Aquaria (EAZA) across several zoos in Europe, and from a single wild population in

Costa Rica (Figure 1). Adults from captive populations (total n=124) were caught during routine veterinary inspections and blood samples were taken either from the brachial or the jugular vein and stored in 90% ethanol or in Queens's lysis buffer (Seutin et al., 1991); feathers from 17 wild nests were collected opportunistically by the MRN field teams in Costa Rica during the 2021 breeding season either from the ground below a nest tree, inside a nest cavity or directly from a chick (n=22). Blood samples from wild birds were taken from the brachial vein of fledglings monitored during the 2021 breeding season (n=3). A single tissue sample was collected from a dead adult in Costa Rica found under a collapsed nest tree. Feather samples were stored dry in separate Ziploc bags. Sample collection was approved by the School of Anthropology and Conservation's Research Ethics Committee at the University of Kent. Samples were transported from Costa Rica to University of Kent under CITES Export Permit 2022-CR5783/SJ (#S8764) and CITES Import Permits 613768/01 and 613768/02, and UK Animal and Plant Health Agency Import Permit ITIMP22.0104.

DNA extraction and amplification

Whole genomic DNA was extracted from blood using an ammonium acetate precipitation method (Nicholls et al.. 2000). DNA was extracted from tissue and feather samples using the DNeasy Blood & Tissue Kit following the protocol as instructed by the manufacturer (Qiagen, UK). Extracted DNA was visualized on agarose gels stained with SYBR-Safe (Thermofisher Scientific Denver, USA) to visually check for DNA quantity and quality. DNA concentration was estimated using a Nanodrop 8000 (Thermo Scientific, Denver, USA.) and all DNA samples were diluted with ddH_20 to a standard concentration of 20ng/µL.

Primer optimization

A set of 24 microsatellite loci were developed by the NERC Environmental Omics Facility (NEOF) at the University of Sheffield, UK, using two blood samples (ID# C19623 and C19624) provided by Chester Zoo, UK. Optimization of loci for Polymerase Chain Reaction (PCR) amplification was initially performed using DNA samples from four individuals. Forward primers of each microsatellite locus were fluorolabelled with either FAM, VIC or NED fluorescent dyes, and loci were arranged in multiplexes defined using Multiplex Manager v1.2 (Holleley & Geerts, 2009). PCRs were performed using a C1000 thermocycler (BioRad, UK) and performed in 10 µL volume reactions, each containing 1.5 µL of extracted DNA at a concentration of 20 ng/ µL, 5 µL of QIAGEN Multiplex PCR Master Mix (Qiagen, UK), 2.5 µL of water and 1 µL of primer multiplex solution at a concentration of 2 µM for each primer. Primer annealing temperatures ranged from 58 to 62 °C, therefore temperature gradient PCRs were performed at 57-63 °C to identify the optimal annealing temperature for the multiplexes. Thermocycler conditions consisted of 15 minutes at 95°C, followed by 25 cycles of 30 seconds at 95°C, 57°C for 90 seconds and 72°C for 60 seconds, followed by a final extension step of 60°C for 30 minutes. PCR products were visualized on a 1% agarose gel alongside a negative control to confirm amplification for each multiplex set. Loci were then amplified across 10 additional samples to test for polymorphism. Loci with more than two alleles and which amplified cleanly were then used to genotype the remaining sample set. To identify individual genotypes, 5 µL of 1:50 dilution of each PCR product was analyzed on an Applied Biosystems 3730xl DNA Analyzer using Big Dye Terminator v3.1 Cycle Sequencing chemistry. Allele scoring was performed using the microsatellite plugin in GENIOUS Prime 2022.1.1.

Microsatellite characterization

We tested for microsatellite null allele frequency, and calculated number of alleles and observed and expected heterozygosity using Cervus 3.0.7 (Marshall et al., 1998) and tested for Hardy-Weinberg Equilibrium (HWE) and linkage disequilibrium (LD) using GenePop on The Web (Raymond & Rousset, 1995; Rousset, 2008) using 24 unrelated individuals selected from the MRN captive population. Loci were tested for sex-linkage by genotyping individuals of known sex with their known sexed offspring, and for loci that suggested sex linkage we blasted the full microsatellite-containing DNA sequence to identify the chromosomal location of the fragment. To check for genotyping error and allelic dropout, we re-extracted and re-amplified a randomly selected 30% of all samples.

Descriptive statistics and inbreeding

We used CERVUS to test for presence of null alleles across the full genotype data set. H_o, H_e and A_r were calculated for each population using the *adegent* (Jombart et al., 2018) and *hierfstat* (Goudet, 2005) packages in R (R Core Team 2022). In studies of small populations, inbreeding may be undetectable using F_{IS}, so we used the program Coancestry 1.0.1.7 to calculate estimated inbreeding coefficients for each individual based on allelic frequencies from genotypic data. Coancestry calculates four inbreeding estimates (Ritland, LynchRt, TrioML, DyadML); we used the triadic maximum likelihood (TrioML) estimate because it allows for inbreeding and has been found to be the estimator that most closely correlates with true relatedness (Hogg et al., 2019; de Jager et al. 2020; Karamanlidis et al., 2021). We calculated mean values of these measures of genetic diversity for each population, tested for significant differences between them using ANOVAs or Kruskal Wallis tests depending on

the distribution of the data and performed appropriate *post-hoc* pairwise tests for differentiation.

Sampled populations

We sampled four captive populations. The precise geographic origin of individuals sampled from MRN and Zooaves (Costa Rica) and Parque de la Conservacion (Colombia) are unknown within their respective countries because they have been sourced via government seizures of illegally traded birds from the pet trade, and therefore no detailed records exist regarding their provenance (individuals were assumed to originate from within the country where they were seized). European samples were sourced from the EEP collection although the provenance of these individuals is unknown. Additionally, we included samples from a wild population of Great Green Macaws collected in Puerto Viejo de Sarapiqui, a core breeding area in Costa Rica monitored by MRN. The full sample set therefore represented five populations that we identified according to their place of origin: MRN, Zooaves, Europe and Colombia; these populations exist as single closed units (i.e. no active transfer of individuals between populations).

Population structure

We performed a Discriminant Analysis of Principal Components (DAPC; Jombart, et al., 2010) using the *k*-mean algorithm and grouped samples according to the five sampled populations. We first used the function *find.clusters* to investigate the number of clusters suggested by the *k*-mean clustering algorithm, which sequentially runs assuming increasing numbers of K to produce clustering solutions which are then compared using a Bayesian Information Criterion approach. We retained all Principal Components, resulting in K=5 (see Supplementary Information, Figure S1a-b). Next, we retained 11 Principal Components as

354 determined by the *a- score*, which relies on repeating the DAPC analysis using randomized
355 groups and then computing a-scores for each group as well as the average a-score; this
356 approach enables an evaluation of the trade-off between power of discrimination and over-
357 fitting (Jombart & Collins, 2015; Jombart et al., 2012). Finally, we repeated the DAPC
358 analysis specifying these clusters and membership probabilities and retained four
359 discriminant functions. In the resulting analysis, a sample was assigned to a cluster when Q>
360 0.7.

361 We used STRUCTURE (Pritchard et al.. 2000), a non-spatial Bayesian clustering method, to
362 examine population structure between all five populations, and within both the Costa Rican
363 and European populations. The Monte-Carlo Markov chain parameters specified for
364 STRUCTURE were: 7 independent simulations, 1,000,000 iterations with a burn-in of
365 200,000 for a range of K values from K=6 to K=8. We implemented the method described
366 by Evanno et al., 2005 to determine the most likely number of clusters by using STRUCTRE
367 HARVESTER (Earl & vonHoldt, 2012).

368 STRUCTURE and DAPC have distinct approaches to identifying patterns of structure within
369 genetic data. DAPC identifies clusters of genetically similar individuals without considering
370 HWE and LD, so that the clusters are identified solely on allelic composition, whereas
371 STRUCTURE assigns individuals to genetic clusters in such a way that *within* assigned
372 population clusters, loci are in HWE and LE (Pritchard et al., 2000). STRUCTURE uses
373 specific population models and relies on several assumptions tan can be difficult to meet and
374 verify, while DAPC focuses on uncovering genetic differences without any previous
375 assumptions by summarizing genetic differentiation between groups and maximizing the

difference between groups while overlooking within-group variation. We chose to use both methods to evaluate genetic structure with and without population assumptions.

Captive breeding genetic management

We used the genotypes produced by the 16 microsatellites to calculate genetic-based R estimates using the program COANCESTRY v. 1.0.1.9 (Wang, 2011) to determine relatedness between current breeding pairs and the relatedness between all individuals of the MRN-Zooaves captive population. We excluded from this analysis the individuals from the MRN population that are going to be released, as they will exit the captive population in the near future as released individuals. COANCESTRY calculates seven different estimators of relatedness, all of which have different assumptions and methodologies, so we used the simulation module of the program to determine the best performing estimator for our data set. These simulations were conducted using allele frequencies obtained for the Great Green Macaw from the microsatellite loci and applying the program Genepop on The Web (Raymond & Rousset, 1995; Rousset, 2008); error rates and missing values were extracted from the genotyped data set, with the settings adjusted to account for inbreeding, using 100 reference individuals and bootstrapping samples. We conducted a Pearson´s correlation between the seven estimators (see Supplementary Table 1 S1) and the "true" relatedness, and selected the triadic likelihood approach, TrioML, as it was the estimator with the highest correlation coefficient that also considers inbreeding.

Genetics of the release population

We calculated H_o, H_e and A_r for the captive and release populations using the *adegent* (Jombart, 2018) and *hierfstat* (Goudet, 2005) packages in R (R Core Team 2022), to understand the genetic diversity to be retained and translocated in the Costa Rican captive

and release populations. For this part of the analysis, we considered the captive population as the population comprised by the MRN individuals that were not going to be released and the entire Zooaves captive population. The statistical significance between the groups was evaluated using a T test or a Wilcoxon test depending on the distribution of the data.

To determine the proportion of alleles that are being transferred from the captive population with the 27 individuals in the reintroduction effort made by MRN, we conducted a simulation in R following Bristol et al., 2013. The model randomly selects different number of individuals taken from the total MRN captive population (3-69) and runs 1000 replicates for different numbers of potentially released individuals, to determine the proportion of alleles potentially captured depending on the number of individuals randomly chosen out of the 69 possible in the source population.

Additionally, we used COANCESTRY to calculate TrioML the relatedness estimate between all members of the release population, as this was the relatedness estimator that had the best fit in the simulation.

RESULTS

Primer optimization

From the 149 samples available for this study, 138 yielded DNA of sufficient quality for PCR amplification of scorable genotypes. Of the 24 microsatellite loci assembled into seven multiplexes and tested on 24 unrelated individuals from MRN (see Table 1), loci Aamb_25306 and Aamb_13122 failed to amplify, and Aamb_11776 and Aamb_22327 amplified with multiple stutter bands that prevented alleles being reliably scored, therefore

these four loci were excluded from further analyses. Locus Aamb_cons_gr3_2 produced an amplicon with reliable scoring, however it only revealed two alleles in the subset of test samples and was therefore not included in further analysis.

Of the remaining 18 loci, none were in linkage disequilibrium. Locus Aamb_24939 was monomorphic, and Aamb_3822 had a high percentage of null alleles (F(Null)>0.2) and was found to be out of Hardy Weinberg Equilibrium (p=0.0078) due to an excess of homozygotes, so this locus was also excluded from further analysis.

We found that genotypes from known females for locus Aamb_3822 did not present expected genotype proportions according to Mendelian laws of inheritance, an indication that this locus may be located in the Z sex chromosome; a BLAST search of the microsatellite sequence for this locus revealed a 94% identity to the blue-crowned parakeet *(Thectocercus acuticaudatus)* CHD1W gene intron E, and a 92% identity to the maroon-bellied parakeet *(Pyrrhura frontalis*) CHD1W gene intron E, which suggests this locus is located on the Z sex chromosome in the Great Green Macaw. Therefore, this locus was also excluded from the final set. Following these tests, a total of 16 loci (Aamb_25191, Aamb_3648, Aamb_12149, Aamb_11776, Aamb_13758, Aamb_24991, Aamb_25928, Aamb_9690, Aamb_9238, Aamb_15058, Aamb_21942, Aamb_18911, Aamb_15017, Aamb_4366, Aamb_5826, Aamb_7958) were observed to be polymorphic with more than two alleles and were therefore used for further analysis. No evidence of allelic dropout was detected in the 30% of samples for which genotyping was repeated.

Levels of genetic diversity and inbreeding

Overall, mean observed heterozygosity (H_o) was 0.613 ± SD 0.17 and expected heterozygosity (H_e) was 0.559 ± SD 0.15 for the global population (Table 2). We found no

444 significant differences in levels of H_0, H_e or Ar between the 5 populations ($p>0.1$, Figures
445 2.a, 2.b and 2c). The highest number of private alleles were detected in the MRN captive
446 population followed by the Colombia captive population which had a substantially smaller
447 (60%) sample size. There was a significant difference in inbreeding coefficient between
448 populations ($p=0.04$, Figure 2d) with latter significant values for paired comparisons in the
449 one tailed Dunn *post hoc* test between Europe and Colombia (p-adj=0.02).

450 ***Population structure: large scale patterns***

451 The STRUCTURE analysis that considered all five populations together indicated the highest
452 ΔK was achieved at K=3 clusters (Figure 3a, Supplementary Information Figure S2 a,b).
453 Using probability assignments estimated by the software, the clusters clearly delineate the
454 Costa Rican, European and Colombian groups of samples. Considering individual
455 assignment when Q>70, all but one of the MRN samples was assigned to cluster 2, with a
456 single sample assigned to cluster 1, all the Zooaves samples belonged to cluster 2, while the
457 wild samples were spread between clusters. With the wild group of samples being the most
458 spread out between clusters, most of the Costa Rican samples (92%) belonged to cluster 2.
459 All but one of the European samples were assigned to cluster 3, with a single sample admixed
460 between the three clusters, and all but one of the Colombian samples were assigned to cluster
461 1, with a single sample also admixed between the three clusters.
462 The DAPC performed with the *k-means* algorithm using all samples suggested the presence
463 of five clusters, with 11 Principal Components explaining 61% of the total variation, and four
464 retained eigen values (Figure 4 a). With this method, the Colombian and EEP populations
465 clustered in different groups with no overlap, while the EEP did present low overlap with the
466 cluster composed in its majority by samples of Costa Rican origin.

Population structure: fine scale patterns

For the STRUCTURE analysis considering only the Costa Rican samples, we obtained the highest ΔK at K=3 clusters (Figure 3 b, Supplementary Figure S3 a,b). We found that samples from wild and captive breeding populations were assigned to all three genetic clusters or were admixed, so their assignment to clusters did not correspond to our a priori populations (Figure 5). The 69 samples belonging to the MRN population were spread out between the clusters, with most individuals belonging to cluster 2 the least to cluster 1, while of the 20 samples belonging to Zooaves, most of the samples belong to cluster 2 or were admixed. Finally, most of the 13 wild samples belong to cluster 1 or were admixed, with no samples belonging to cluster 3. The wild sample that was assigned to cluster 3 had >93% of the genotype missing, and two of the samples assigned to cluster 3 had 68% and 87% of the genotype missing, so we do not consider these assignments as relevant to determine population of origin. A single Costa Rican sample was assigned to the European cluster, with a membership probability of 75% with a genotype missing only 18%, which may indicate this sample has a similar genetic composition to those individuals represented in the sampled European collections.

Finally, the STRUCTURE analysis performed on the European samples also revealed the highest ΔK to be K=3 clusters (Figure 3.c, Supplementary Figure S4 a,b); within this group of samples, 40% belonged to group 1, 31% to cluster 3, 13% to cluster 2 and 13% were admixed between the three groups.

The DAPC method split the Costa Rican samples into clusters 1, 2 and 4 and EEP samples into clusters 1 and 3. In the Costa Rican sample set, most of the captive samples were assigned to cluster 1, and the only captive sample that didn't belong to this cluster was

assigned as the sole representative of cluster 2; on the other hand, wild samples were assigned with Costa Rican captive samples to cluster 1, but importantly, they also formed an exclusive group, cluster 4. Cluster 3 holds most of the European samples, with the exception of 3 individuals that were assigned with the Costa Rican samples in cluster 1.

Captive breeding genetic management

We obtained Relatedness estimators for the nine breeding pairs active in the MRN breeding center, and our results show variation in relatedness amongst the breeding pairs, ranging from 0 to 0.264 (Table 3). Our results suggest that the assumed relatedness of 0 was true for three breeding pairs: in all these breeding pairs both individuals come from private donors or unknown origin and have unknown geographic origin. For breeding pair 2, the relatedness coefficient is close to an R value of 0.25 which is expected between half sibling relationships or half siblings/ avuncular/grandparent–grandchild. Both members of this pair were born in captivity, but their records are incomplete; for one of them there is no information available about their ancestry and the other was born to founders of unknown relation. Breeding pairs 3, 5, 6 and 8 have a relatedness coefficient closest to the expected R of 0.03 between second cousin relations. For Breeding pair 3, one member of the pair was brought to the breeding population by the environmental authorities with unknown history or place of origin, while the partner was born in captivity with unknown date of birth and ancestry. In contrast, both members of Breeding pairs 5 and 8 are second generation individuals born in captivity to founders of unknown relations, while both members of Breeding pair 6 have unknown origins and relationships. Finally, both members of Breeding pair 9 were born in captivity, but for one of them there is no record of their ancestry whilst the partner was the progeny of founders of unknown relationship.

We further evaluated the relatedness between the whole Costa Rican captive population considered (MRN and Zooaves) and obtained a Relatedness plot with the TrioML estimated values between all possible pairs of individuals (Figure 6).

Genetics of the release population

In relation to the genetic diversity of the captive and release populations, we found that H_e, Relatedness and Ar were higher in the population that will remain in captivity, (Table 4), but only the Relatedness was statistically significant (p-val= 0.03). We also found that the proportion of alleles that will be transferred in the 27 released individuals represents 70% of the total allelic diversity present within the MRN captive population (Figure 7). However, the proportion of alleles that could be transferred by 27 individuals randomly selected by the model is somewhat higher than this (median value=75%).

Finally, amongst the individuals that will comprise the release population we detected relatedness values in the orders of full siblings and cousins (Figure 8): there are seven groups of siblings due to be released in this population, and as well as two groups of cousins, which is not surprising given that these individuals destined for release all come from 12 breeding pairs.

DISCUSSION

Our study presents for the first time an insight into levels of population genetic diversity in the Great Green Macaw, and how that variation is distributed among wild and captive populations in Costa Rica as well as across captive populations in Colombia and the EEP.

Comparison of levels of genetic diversity

Levels of genetic diversity did not differ significantly between the different populations, except between the European and Colombian populations. The wild Costa Rican population had similar levels of genetic diversity when compared to the two Costa Rican captive populations, however it is worth noting that the wild population has the smallest sample size and the most incomplete genotypes due to failure in the amplification of the full set of microsatellite loci because of lower quality DNA extracted from shed feathers (54% amplification success in feather samples, in contrast to a 94% amplification success in samples with DNA extracted from blood).

Based upon a mean wild H_e, Witzenberger & Hochkirch (2011) suggest that captive populations should have levels of $H_e = 0.54$ to maintain 90% of the natural genetic variation. We found that mean H_e in our captive populations ranged between 0.517 and 0.622, which indicates an acceptable level of H_e retained in the *ex-situ* populations. Levels of genetic diversity among the Great Green Macaws were found to be comparable to those of other macaw species. The Spix Macaw (EW), Lears Macaw (EN) (Presti, et al., 2011), Red fronted macaw (CR) (Blanco et al., 2021) have a H_o that range between 0.48 and 0.63, while the Hyacinth Macaw (VU) has the lowest H_o= 0.32, the Blue throated macaw (CR) has a higher H_o= 0.68 (Campos et al., 2021) with the Scarlet Macaw (LC) H_o= 0.86 has the highest (Escalante-Pliego et al., 2022). H_o was also comparable to other parrot species, like the Bahama parrot (H_o=0.69), the South African parrot (H_o=0.581), Blue fronted parrot (H_o=0.869), Cuban amazon (H_o=0.64 – 0.77), the Swift parrot (Ho = 0.679), the Ring-necked parakeet (H_o= 0.662) and the Kakapo (H_o = 0.489). In these studies (Leite et al., 2008; Russello et al., 2010; Bergner et al., 2014; Stojanovic et al., 2018; Coetzer et al., 2020) , H_o

ranged between 0.32 to 0.869, and our findings both per population and globally fall within this range. Most of these studies have a low sample size associated with the challenges of obtaining samples from parrots, so comparations must be made with caution.

Fine scale patterns of genetic structure

The two captive Costa Rican populations have very similar genetic composition and although the STRUCTURE and DAPC analyses indicate that the wild population appears to be a distinct cluster from them, it is more closely associated with these captive clusters than any other sampled population. One interpretation is that genetic drift, which can arise from founder effects or generations of population isolation, has not substantially altered the genetic composition of these two captive populations compared to that of the wild population. The two captive populations were established relatively recently from individuals that were seized from pet trade by Costa Rican government agencies and are therefore most likely to have been sourced from wild populations within that country.

STRUCTURE did not assign any individuals sampled from the European captive collection to a cluster other than to itself, suggesting that the microsatellite marker set was not able to confidently assign geographic origin to any of the European captive samples. The DAPC approach assigned three individuals from the European population to cluster 1, which represents wild and captive Costa Rican samples, suggesting their origin in this country. The genetic representation of Costa Rica birds in the European group, but no similar clustering of individuals from the Colombian group, suggests there are unlikely to be any Colombian-sourced birds amongst the sampled European collection.

Intriguingly, our STRUCTURE analysis of the sampled European captive population detected the presence of three genetic clusters within it, however in the absence of any

information on the provenance or origins of the genotyped individuals it is not possible to infer much beyond the probability that the European captive population has been established using founders from at least three different geographic sources. Furthermore, the EEP has a population of about 150 individuals so our study does not include the whole population managed by the EEP, therefore there may be genetic diversity present in this international captive breeding program that has yet to be described; this calls for further efforts to more comprehensively sample and genotype the European population, so that more complete information is available to guide breeding and management decisions.

Large scale patterns of genetic structure

Our Structure and DAPC analyses both suggest the existence of identifiably different clusters across Costa Rica, Colombia, and Europe. The Colombian and European captive populations were assigned to distinct groups with little overlap, indicating substantial genetic differentiation between these populations. The extent of differentiation between Colombian and Costa Rican populations could be explained by geographical separation and may therefore reflect a lack of gene flow associated with isolation by distance, which considers genetic differentiation and structure as a function of Euclidean distance (Wright, 1943). This evolutionary process, whereby genetic differences between individuals and populations increase with geographic distance, is based on the assumption of limited dispersal that leads to restricted mating (Sánchez-Ramirez et al., 2018). Great Green Macaws occur in Panama with a continuous distribution extending into Colombia, but there is a gap in its distribution from Panama to northern Costa Rica, which could further contribute to genetic differentiation between populations in Costa Rica and populations in Panama and Colombia. Spatially isolated populations can have high levels of genetic diversity due to the accumulation of

genetic differences (Taylor et al., 2021) associated with local adaptation in response to geographically variable selection (Tiffin & Ross-Ibarra, 2014). Since we don't have information about the specific geographic origin of the samples, we are restrained by our data set and are unable to run a test to detect Isolation by Distance, but this result might represent a first indication that this phenomenon is playing a role in shaping genetic structure in the species.

Information on extent of genetic structure and differentiation can be valuable for landscape-level population management when determining conservation management units; if such units are based on demographic independence supported by genetic data, then they may reflect true population differentiation that may need to be preserved by delimitation (Coates et al., 2018; Keller et al., 2015; Manel et al., 2003). The Great Green Macaw populations in Costa Rica and Colombia might therefore be considered sufficiently different to warrant being managed separately, however further geographic sampling across this widespread species' range will be necessary before more definitive conclusions can be made.

Genetic differentiation was consistently higher between the Colombian population and the European / Costa Rican populations, indicating that the European population that we sampled has a genetic make-up closer to the Costa Rican populations than to the Colombian population. Some individuals from the European collection were assigned to the same cluster as individuals from known Costa Rican origin, which suggests those individuals are likely to have their origins within this country. The remaining 19 European samples of unknown origin have a very different genetic makeup to the captive population sampled from Colombia, suggesting that Colombian genetic diversity captured in our study has no representation in the European population for which samples were available.

Management implications

Captive breeding success is strongly determined by genetic processes such as loss of genetic diversity through genetic drift and accumulation of inbreeding (Willoughby et al., 2015). Therefore, management strategies should be implemented to minimize inbreeding, reduce problems associated with inbreeding depression and risks of genetic adaptation to captivity (Witzenberger & Hochkirch 2013). The results obtained in this research aimed to address the three main objectives of captive populations: (i) to maintain captive populations that are genetically representative of wild populations, (ii) to maintain and maximize retention of genetic diversity over time, and (iii) to provide individuals to establish viable reintroduced populations (Frankham et al., 2002, IUCN/SSC 2013); thus, they can be applied to captive management to inform conservation planning and management decisions for the Great Green Macaw in an improved, galvanized and research based way.

Ex-situ/In-situ diversity

Many studies compare the genetic diversity in wild and captive populations to determine how much genetic diversity is being preserved in *ex situ* conservation programs (Ramirez, et al., 2006; West et al., 2018, Kleinman-Ruiz et al., 2019, Morrison et al., 2020, White et al., 2022) and use it to identify the geographical origin of captive representatives (Pasachnik, et al., 2020; Oklander et al., 2009; Ogden & Linacre 2015). The genetic diversity of the captive populations sampled in Costa Rica appears to be, to a large degree, representative of that found in the wild population that was sampled. Our results are therefore, a first step in ensuring that the captive breeding program of the Great Green Macaw in Costa Rica might be a suitable source for the selection of individuals for reintroduction.

649 The majority of the captive samples in this study originate from unidentified locations, and
650 the lack of widespread sampling across the different wild populations hinders our ability to
651 fully interpret our results in this context, but we can ascertain that all of the alleles present in
652 the monitored population in Sarapiqui, Costa Rica, are present in captivity. Understanding
653 how genetic diversity is geographically distributed would allow a better understanding that
654 could enable more tailored decision-making for captive management. For the Costa Rican
655 population, current efforts to sample wild populations are ongoing to further understand the
656 population genetics status of the species in the wild; these efforts will strengthen a joint "One
657 Plan approach" as recommended by the IUCN (Redford et al. 2012).

658 The European captive population would further benefit from large scale, species range
659 sampling to understand its representativeness of the genetic diversity of the species, as well
660 as genotyping of the whole population. Given that the EEP captive individuals have an
661 unknown geographical origin, that this ex-situ population could contain genetic
662 representation of multiple subpopulations, sub-species, and accidental hybrids of the two, so
663 relating captive diversity to wild counterparts would be valuable in terms of management to
664 comply with *ex situ* genetic goals, but also for mating pair determination considering the
665 possibility of having individuals representing more than one subspecies.

666 *Genetic management of captive breeding populations*

667 Integrating genetic results into management remains a challenge because population
668 managers might lack the expertise to develop such research, and researchers might be
669 unfamiliar with how to turn genetic data into a conservation tool for managers (Normal, et
670 al., 2018). Our research enables a step towards bridging the gap between conservation
671 genetics and captive breeding management for the Great Green Macaw, by providing genetic

672 data that can be used to understand the relatedness between founders and for future breeding
673 pair recommendations based on molecular data.

674 From examining the ancestry of the current active breeding pairs, we identified that most
675 individuals were either born in captivity to parents of unknown origin and relations, or have
676 unknown origin themselves, and that importantly, some of the original founders that
677 produced members of the current breeding pairs are not part of the current MRN captive
678 population, and no samples were stored. Molecular data like those produced in this study can
679 commonly be used to infer unknown parentage in a captive population (Ivy et al., 2009 ;
680 Ferrie et al., 2013; Miller-Butterworth et al., 2021; Weng et al., 2021), but we could not
681 reconstruct these relationships because samples of all potential parents were not available for
682 genetic testing. In such instances, molecular markers can be used to estimate their pairwise
683 relatedness with other living members of the captive population (Ivy et al., 2009), as we have
684 done.

685 Variance in relatedness coefficient between breeding pairs might be explained by the
686 potential high levels of relatedness between individuals in a genetically depauperate
687 population (Bergner et al., 2014; Hogg 2019) or perhaps by unintentional pairing between
688 related individuals due to incomplete knowledge of individual ancestry.

689 Without a pedigree available for the Great Green Macaw, an alternative approach is to use
690 genetic-based estimates of pairwise relatedness to inform pairing decisions. New pairings
691 could be formed based on achieving lower relatedness coefficient between a breeding pair
692 and conversely, some pairings having high relatedness values should be avoided. For the 9
693 breeding pairs considered in MRN, potential pairings with other members of the MRN
694 captive population resulted in reduced relatedness coefficients (see Figure 6), and

695 furthermore, some pairings involving Zooaves-MRN birds resulted in low relatedness
696 coefficients; these pairings could be explored by both organizations to exchange birds
697 between the flocks or to rearrange current breeders within MRN.

698 However, current best practice for making breeding pair recommendations in captive
699 breeding programs is based on the use of multigenerational pedigree data to estimate kinship
700 between living individuals within a captive population (Galla et al., 2020). The kinship
701 between two individuals is the probability of two alleles at a given locus randomly drawn
702 from each individual being identical by descent from a common ancestor (Falconer 1981).
703 The use of pedigree data is targeted at meeting demographic and genetic goals for the
704 population and aims to minimize the population average kinship by breeding
705 underrepresented individuals with low individual mean kinship (Ivy et al., 2009). This
706 approach maximizes founder representation and minimizes inbreeding over time
707 (Willoughby et al. 2015).

708 With the results of this study, (i) the empirical relatedness between founders could be
709 estimated and thus, the assumption of all founders being unrelated could be bypassed, and
710 (ii) the relatedness matrix between the rest of the population could also be estimated, thus
711 providing information that can be used to avoid erroneous estimates of mean kinship and
712 inbreeding coefficients for optimized management decisions (Russello & Amato, 2004).
713 Thus, the genetic data can be integrated alongside the known family relationships within the
714 captive population(s) to form a pedigree for conservation management the Great Green
715 Macaw in Costa Rica. Empirical estimates of relatedness can create a baseline of known
716 relatedness that can be integrated into traditional pedigree management; this approach would
717 represent a further recommended use of the genetic relatedness estimators since breeding pair

718 recommendations based on pedigree incorporate additional elements other than relatedness
719 for the pair. For instance, the pedigree management program, PMx (Lacy et al., 2012), allows
720 for a set up in which a matrix of empirical relatedness is considered for the determination of
721 an empirical metric of kinship to create breeding pair recommendations based on the Mate
722 Suitability Index (MSI) and Mean Kinship (MI).

723 Presently, MRN and Zooaves captive populations are not being genetically managed using a
724 pedigree approach. We recommend the construction of a pedigree that involves both *ex situ*
725 populations for joint management and the integration of these results in a combined approach
726 using realized relatedness that augments data in the pedigree. We believe that the integration
727 of MRN and Zooaves flock as a joint management unit, the construction of a pedigree and
728 the introduction of genetic management using PMx with empirical relatedness values, could
729 help optimize the captive breeding program for the Critically Endangered Great Green
730 Macaw.

731 *Genetic perspective on future planned reintroductions*

732 Future plans to reintroduce Great Green Macaws into different regions of Costa Rica where
733 there are suitably sized areas of high-quality and restored habitat form an important
734 component of the overall conservation interventions for recovering this critically-endangered
735 species. MRN has a selected population of 27 captive born individuals that are going to be
736 released in 2024, and the genetic makeup of this subset of individuals is now known because
737 they were all included in this study.

738 Our genetic analysis revealed that H_e, Ar and Relatedness are higher in the captive population
739 than in the release population, but the difference is only statistically significant for the
740 relatedness. Higher relatedness in the captive population can be explained by two facts: 19

MRN captive born individuals will remain in captivity, and in this group there are pairs of siblings and cousins, additionally 8 of this 19 captive born birds have no records of ancestry so they could also be related among themselves and other members of the group; on the other hand, we don't have the ancestry records of the Zooaves flock, and since they are a captive breeding center, we can expect higher levels of relatedness among this group due to successful breeding.

Our simulation of selection of individuals for reintroduction indicates that MRN has the potential to release more allelic diversity available in the whole captive population if they select individuals randomly compared to the specific 27 individuals earmarked for reintroduction. However, criteria for selecting individuals for release needs to consider factors beyond genetics, such as breeding potential, veterinary health, and social behavior. The proportion of the captive population intended for release is made up entirely of individuals born in different cohorts in captivity since 2017, and they only partly represent the living reproductive output of the breeding population. They are thus a subset of the allelic diversity of the MRN population that is associated with those breeding pairs that have bred successfully since the captive breeding project was initiated. Their allelic representation of close to three quarters of the allelic diversity of the whole captive population is high, considering they come from 12 breeding pairs.

The practicalities of selecting individuals for release often mean that relatedness and allelic diversity need to be considered alongside a range of more pragmatic issues, and managers relinquish control – upon release – of which individuals form breeding pairs to produce F1 offspring in the reintroduced population. Nonetheless, we note that some possible pair combinations could produce offspring with high levels of relatedness in which there is a high

764 proportion of shared alleles that are identical by descent (IBD) in this population; such
765 instances could contribute to heightened levels of inbreeding and subsequent reduced long-
766 term viability in the population that becomes established following the reintroduction
767 (Groombridge et al., 2012). It is now recognized that breeding between related individuals
768 results in inbreeding depression, which affects population growth and persistence over time
769 (Frankham, 2005). Inbreeding depression is of particular concern in small populations, in
770 which inbreeding can act in combination with genetic drift to further cause genetic diversity
771 loss (Frankham et al. 2002). The direct risk of mating between related individuals lies in the
772 increased risk of homozygosity in deleterious recessive alleles and the increased genome
773 wide homozygosity of IBD alleles generated by consanguineous mattings (Townsend &
774 Jamieson, 2013). Regarding the forthcoming release of Great Green Macaws, there are many
775 potential first order and second order relationships that could potentially produce highly
776 related and inbred wild born individuals, this risk is exacerbated by the small size of the
777 population. These two considerations pose a threat for inbreeding depression in the future
778 when wild breeding starts.

779 MRN management decisions for the location of the release populations have not been
780 finalized. To our knowledge there are two potential strategies: The release population can be
781 released either into an area where the species was known to previously occur but from where
782 it was extirpated in recent years, or into an area of proximity to the existing wild breeding
783 population.

784 If the intention is for the released population to form a new, independent founding population
785 for the species in the extirpated area, this independent population will only be supplemented
786 by new cohorts of captive bred individuals, most likely descendants of the same successful

787 captive breeding pairs that originated from the first release cohort. This scenario would entail
788 a relatively limited pool of individuals for pairing opportunities and so could lead to greater
789 levels of inbreeding and inbreeding depression which could threaten the long-term
790 establishment of the reintroduced population. If the selection for future breeding pairs for
791 Zooaves and MRN is done considering the joint flock and genetic recommendations, the next
792 cohorts to be released could have a lower estimated relatedness among themselves and this
793 population, potentially lowering the future inbreeding risk. Additional risks of releasing the
794 27 individuals as an independent founding population are those generally associated with low
795 population density and small populations, which include stochastic population loss and the
796 Allee effect (Lande 1988, Courchamp et al., 1999) and genetic drift (Masel, 2011; Keller et
797 al., 2012)

798 Conversely, if the individuals are released in an area where they will supplement an existing
799 population, then there is a greater chance for the released individuals to form pairings with
800 individuals from the wild, thus the probability of consanguineous matings might be lower.
801 Additionally, releasing new captive individuals into the wild populations might represent a
802 source of geneflow. This evolutionary force can introduce new alleles into the wild
803 population aiding to mask the expression of deleterious ones and boosting the population´s
804 adaptative potential (Willi et al. 2022), and for inbred populations, such translocations can
805 alleviate genetic load, inbreeding depression and reduced genetic variation (Weeks et al.,
806 2011). On the other hand, mixing genetically divergent stock can lead to outbreeding
807 depression, in which the fitness of the population is reduced due to genetic incompatibilities,
808 the disruption of local adaptations and the influx of poorly adapted genes from divergent
809 environmental conditions (Frankham et al., 2011).

In the face of gaps in knowledge about the full genetic relationships amongst different sub-populations of an endangered species, balancing these risks is inevitably a challenge; indeed, weighing up these risks between management options can commonly lead to inaction (Weeks et al., 2011). However, the detrimental effects of outbreeding may have a less negative effect on long term population viability compared to the short-term risks associated with inbreeding and loss of genetic diversity via drift. In the case of the Great Green Macaw, we now have a somewhat clearer understanding of the genetics of the captive breeding populations, but we are still beginning to understand that of the wild populations. The genetic or demographic effects of either management decision are difficult to predict, especially because we have limited knowledge of the distribution and type of genetic diversity in the wild. We therefore advocate for continued and more extensive genetic monitoring to ascertain demographic and genetic effects to aid management decision for this species.

Genetic management: Genetic rescue

Introducing novel diversity from elsewhere in the species' geographic range, such as from captive stock in Colombia, which our study has shown to contain substantially different genetic diversity, could potentially bring benefits of 'genetic rescue' (Tallmon et al., 2004) whereby novel genes (alleles) mitigate the effects of inbreeding to increase levels of heterozygosity and hence potentially lead to increased fitness. The desired outcome of this management practice is a demographic response and increase in absolute fitness at a population level to reduce population extinction risk (Whiteley et al., 2015). This approach might mitigate some of the problems of inbreeding depression that are one possible explanation for the low levels of productivity reported for the MRN captive population in Costa Rica (MRN *personal com*). Conversely, introgressing novel genes brings with it a risk

of outbreeding depression if the population source of those genes is too genetically differentiated (Frankham, et al. 2010). The balance of these risks and potential benefits is the reason why there is considerable debate around the use of genetic rescue and why it remains relatively underused and controversial in conservation efforts to recover critically endangered species (Bell et al., 2019). It is currently not possible to interpret which outcome may be the most likely scenario for the Great Green Macaw without further intensive sampling across the species' widespread geographic range and additional captive populations.

CONCLUSIONS

This study has described genetic structure in the Great Green Macaw, that at a large scale corresponds predominantly to place of origin of the samples (i.e. Costa Rica, Colombia and Europe). At a finer scale, the structure found in Costa Rica doesn't correspond to captive and wild populations, which leads us to interpret from these results that genetic drift may not have had time to act on the captive populations to differentiate their genetic diversity between the captive breeding centers, and that there is still similar extent of genetic diversity found in both captive populations; the structure found in the European samples of the EEP led us to hypothesize that the clusters detected might represent three places of origin for the captive samples, although our analyses have not been able to identify phylogeographic origin of those samples. We acknowledge that the interpretation of our results is limited because we don't know the geographic origin of the samples in the captive populations and we have a limited sample size from the widely distributed wild population. However, we believe that this initial description of the genetic diversity and genetic structure of Great Green Macaw can lead the way for further research into the wild populations to (i) strengthen the understanding of

genetic diversity in the wild, (ii) aid in the evaluation of wild genetic representativeness in captive population and (iii) contribute to determining geographic origin for captive samples. We recommend that MRN and Zooaves captive breeding populations in Costa Rica manage their flock jointly, and integrate molecular measures of relatedness and the use of a pedigree in their future pairing decisions. Our study highlights the value of applying DNA marker data and molecular estimates of relatedness in captive breeding and reintroduction strategy, and we urge managers to integrate them into future conservation actions.

REFERENCES

Aguiar, J. D. P., Gomes, P. F. F., Hamoy, I. G., Santos, S. E. B. D., Schneider, H., & Sampaio, I. (2018). Loss of genetic variability in the captive stocks of tambaqui, Colossoma macropomum (Cuvier, 1818), at breeding centres in Brazil, and their divergence from wild populations. *Aquaculture Research*, *49*(5), 1914-1925.

Alcaide, M., Negro, J. J., Serrano, D., Antolín, J. L., Casado, S., & Pomarol, M. (2010). Captive breeding and reintroduction of the lesser kestrel Falco naumanni: a genetic analysis using microsatellites. *Conservation genetics*, *11*, 331-338.

Allendorf, F. W., Luikart, G. H., & Aitken, S. N. (2012). *Conservation and the genetics of populations*. John Wiley & Sons.

Attard, C. R. M., Möller, L. M., Sasaki, M., Hammer, M. P., Bice, C. M., Brauer, C. J., ... & Beheregaray, L. B. (2016). A novel holistic framework for genetic-based captive-breeding and reintroduction programs. *Conservation Biology*, *30*(5), 1060-1069

Ayala-Burbano, P.A., Galetti Junior, P.M., Wormell, D. *et al.* Studbook and molecular analyses for the endangered black-lion-tamarin; an integrative approach for assessing genetic diversity and driving management in captivity. *Sci Rep* 10, 6781 (2020). https://doi.org/10.1038/s41598-020-63542-2

Ballou, J. Calculating inbreeding coefficients from pedigrees. In Genetics and conservation: a reference for managing wild animal and plant populations (ed. Schonewald-Cox, C. M., Chambers,

S. M., MacBryde B, T. W.) (The Benjamin/Cummings Publishing Company Inc., Menlo Park, 1983).

Blanco, G., Morinha, F., Roques, S., Hiraldo, F., Rojas, A., & Tella, J. L. (2021). Fine-scale genetic structure in the critically endangered red-fronted macaw in the absence of geographic and ecological barriers. *Scientific reports*, *11*(1), 556.

Bell, D. A., Robinson, Z. L., Funk, W. C., Fitzpatrick, S. W., Allendorf, F. W., Tallmon, D. A., & Whiteley, A. R. (2019). The exciting potential and remaining uncertainties of genetic rescue. *Trends in Ecology & Evolution*, *34*(12), 1070-1079.

Bergner, L. M., Jamieson, I. G., & Robertson, B. C. (2014). Combining genetic data to identify relatedness among founders in a genetically depauperate parrot, the kakapo (Strigops habroptilus). Conservation Genetics, 15(5), 1013–1020. https://doi.org/10.1007/ s10592-014-0595-

BirdLife International. 2020. Great Green Macaw (Ara ambiguus) - BirdLife species factsheet [Internet]. [accessed 2021 Jan 5]. http://datazone.birdlife.org/species/factsheet/greatgreen-macaw-ara-ambiguus

Bristol, R. M., Tucker, R., Dawson, D. A., Horsburgh, G., Prys-Jones, R. P., Frantz, A. C., ... & Groombridge, J. J. (2013). Comparison of historical bottleneck effects and genetic consequences of re-introduction in a critically endangered island passerine. *Molecular ecology*, *22*(18), 4644-4662.

Brichieri-Colombi, T. A., Lloyd, N. A., McPherson, J. M., & Moehrenschlager, A. (2019). Limited contributions of released animals from zoos to North American conservation translocations. *Conservation Biology*, *33*(1), 33-39.

Campos, C. I., Martinez, M. A., Acosta, D., Diaz-Luque, J. A., Berkunsky, I., Lamberski, N. L., ... & Wright, T. F. (2021). Genetic diversity and population structure of two endangered neotropical parrots inform in situ and ex situ conservation strategies. *Diversity*, *13*(8), 386.

Carroll, E. L., Bruford, M. W., DeWoody, J. A., Leroy, G., Strand, A., Waits, L., & Wang, J. (2018). Genetic and genomic monitoring with minimally invasive sampling methods. *Evolutionary applications*, *11*(7), 1094-1119.

Chassot O, Arias GM, Powell GVN. 2007. Biologia de la Conservacion de Ara ambiguus en Costa Rica, 1994-2006. Mesoamerica. 11(2):43–49.

Coates, D. J., Byrne, M., & Moritz, C. (2018). Genetic diversity and conservation units: dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution*, *6*, 165.

Coetzer, W. G., Downs, C. T., Perrin, M. R., & Willows-Munro, S. (2020). Influence of historical and contemporary habitat changes on the population genetics of the endemic South African parrot Poicephalus robustus. *Bird Conservation International*, *30*(2), 236-259.

Courchamp, F., Clutton-Brock, T., & Grenfell, B. (1999). Inverse density dependence and the Allee effect. *Trends in ecology & evolution*, *14*(10), 405-410.

de Jager, D., Harper, C. K., & Bloomer, P. (2020). Genetic diversity, relatedness and inbreeding of ranched and fragmented Cape buffalo populations in southern Africa. *Plos one*, *15*(8), e0236717.

De Barba, M., Waits, L. P., Garton, E. O., Genovesi, P., Randi, E., Mustoni, A., & Groff, C. (2010). The power of genetic monitoring for studying demography, ecology and genetics of a reintroduced brown bear population. *Molecular ecology*, *19*(18), 3938-3951.

DeWoody, J. A., Harder, A. M., Mathur, S., & Willoughby, J. R. (2021). The long-standing significance of genetic diversity in conservation. *Molecular Ecology*, *30*(17), 4147-4154.

Earl, D. A., & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources*, *4*, 359-361.

Eizaguirre, C., & Baltazar-Soares, M. (2014). Evolutionary conservation—evaluating the adaptive potential of species. *Evolutionary Applications*, *7*(9), 963-967.

Escalante-Pliego, P., Matías-Ferrer, N., Rosas-Escobar, P., Lara-Martínez, G., Sepúlveda-González, K., & Raigoza-Figueras, R. (2022). Genetic Diversity and Population Structure of Mesoamerican Scarlet Macaws in an Ex Situ Breeding Population in Mexico. *Diversity*, *14*(1), 54.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, *14*(8), 2611-2620.

Ewen, J. G., Armstrong, D. P., Parker, K. A., & Seddon, P. J. (Eds.). (2012). *Reintroduction biology: integrating science and management*. John Wiley & Sons.

Falconer DS. 1981. Introduction to quantitative genetics. 2nd ed. New York: Longman

Farquharson, K. A., Hogg, C. J., & Grueber, C. E. (2021). Offspring survival changes over generations of captive breeding. *Nature communications*, *12*(1), 3045.

940 Ferrie, G. M., Cohen, O. R., Schutz, P., Leighty, K. A., Plasse, C., Bettinger, T. L., & Hoffman, E.
941 A. (2013). Identifying parentage using molecular markers: improving accuracy of studbook records
942 for a captive flock of marabou storks (Leptoptilus crumeniferus). *Zoo biology*, *32*(5), 556-564.

943 Fjeldså J, Krabbe N, Ridgely RS. 1987. Great Green Macaw Ara ambigua collected in northwest
944 Ecuador, with taxonomic comments on Ara militaris. Bull Br Ornithol Club. 107:28–31

945 Frantz, A. C., Pope, L. C., Etherington, T. R., Wilson, G. J., & Burke, T. (2010). Using isolation-by-
946 distance-based approaches to assess the barrier effect of linear landscape elements on badger (Meles
947 meles) dispersal. *Molecular Ecology*, *19*(8), 1663-1674.

948 Fisch, K. M., Kozfkay, C. C., Ivy, J. A., Ryder, O. A., & Waples, R. S. (2015). Fish hatchery genetic
949 management techniques: integrating theory with implementation. *North American Journal of*
950 *Aquaculture*, *77*(3), 343-357.

951 Frankham R, Ballou JD, Briscoe DA (2010). Introduction to conservation genetics, 2nd edn.
952 Cambridge University Press, Cambridge

953 Frankham, R., Briscoe, D. A., & Ballou, J. D. (2002). *Introduction to conservation genetics*.
954 Cambridge university press.

955 Frankham R (2005) Genetics and extinction. Biol Cons 126(2):131–140

956 Frankham R, Ballou JD, Eldridge MDB, Lacy RC, Ralls K, Dudash MR, Fenster CB (2011)
957 Predicting the probability of outbreeding depression. Conserv Biol 25(3):465–475

958 Frankham R, Ballou JD, Ralls K, Eldridge M, Dubash MR, Fenster CB, Lacy RC, Sunnucks P (2017)
959 Genetic management of fragmented animal and plant populations. Oxford University Press, Oxford.

960 Galla, S. J., Moraga, R., Brown, L., Cleland, S., Hoeppner, M. P., Maloney, R. F., ... & Steeves, T.
961 E. (2020). A comparison of pedigree, genetic and genomic estimates of relatedness for informing
962 pairing decisions in two critically endangered birds: Implications for conservation breeding
963 programmes worldwide. *Evolutionary Applications*, *13*(5), 991-1008

964 Garnier, S., Alibert, P., Audiot, P., Prieur, B., & Rasplus, J. Y. (2004). Isolation by distance and sharp
965 discontinuities in gene frequencies: Implications for the phylogeography of an alpine insect species,
966 Carabus solieri. Molecular Ecology, 13(7), 1883– 1897. https://doi. org/10.1111/j.1365-

967 Gooley, R. M., Dicks, K. L., Ferrie, G. M., Lacy, R. C., Ballou, J. D., Callicrate, T., ... & Pukazhenthi,
968 B. S. (2022). Applying genomics to metapopulation management in North American insurance

populations of southern sable antelope (Hippotragus niger niger) and addra gazelle (Nanger dama ruficollis). *Global Ecology and Conservation*, *33*, e01969.

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular ecology notes*, *5*(1), 184-186.

Groombridge, J. J., Raisin, C., Bristol, R., & Richardson, D. S. (2012). Genetic consequences of reintroductions and insights from population history. *Reintroduction biology: Integrating science and management*, 395-440.

Hammerly, S.C., Morrow, M.E. & Johnson, J.A. (2013). A comparison of pedigree- and DNA-based measures for identifying inbreeding depression in the critically endangered Attwater's prairie-chicken. Mol. Ecol. 22, 5313–5328. Henkel, J.R., J

Hammerly, S. C., de la Cerda, D. A., Bailey, H., & Johnson, J. A. (2016). A pedigree gone bad: increased offspring survival after using DNA-based relatedness to minimize inbreeding in a captive population. *Animal Conservation*, *19*(3), 296-303.

Henkel, J. R., Jones, K. L., Hereford, S. G., Savoie, M. L., Leibo, S. P., & Howard, J. J. (2012). Integrating microsatellite and pedigree analyses to facilitate the captive management of the endangered Mississippi sandhill crane (Grus canadensis pulla). *Zoo biology*, *31*(3), 322-335.

Hofreiter, M., & Stewart, J. (2009). Ecological change, range fluctuations and population dynamics during the Pleistocene. *Current biology*, *19*(14), R584-R594.

Hogg, C. J., Wright, B., Morris, K. M., Lee, A. V., Ivy, J. A., Grueber, C. E., & Belov, K. (2019). Founder relationships and conservation management: empirical kinships reveal the effect on breeding programmes when founders are assumed to be unrelated. *Animal Conservation*, *22*(4), 348-361.

Holleley, C. E., & Geerts, P. G. (2009). Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. *Biotechniques*, *46*(7), 511-517.

Ivy, J. A., Miller, A., Lacy, R. C., & DeWoody, J. A. (2009). Methods and prospects for using molecular data in captive breeding programs: an empirical example using parma wallabies (Macropus parma). *Journal of Heredity*, *100*(4), 441-454.

IUCN/SSC (2013). Guidelines for Reintroductions and Other Conservation Translocations. Version 1.0. Gland, Switzerland: IUCN Species Survival Commission, viiii + 57 pp

IUCN/SSC (2014). Guidelines on the Use of Ex Situ Management for Species Conservation. Version 2.0. Gland, Switzerland: IUCN Species Survival Commission.

Jombart, T., & Collins, C. (2015). A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0. 0. *London: Imperial College London, MRC Centre for Outbreak Analysis and Modelling*.

Jombart, T., Kamvar, Z. N., Collins, C., Lustrik, R., Beugin, M. P., Knaus, B. J., & Jombart, M. T. (2018). Package 'adegenet'. *Github Repository. Available online: https://github. com/thibautjombart/adegenet (accessed on 11 November 2021)*.

Karamanlidis, A. A., Skrbinšek, T., Amato, G., Dendrinos, P., Gaughran, S., Kasapidis, P., ... & Stronen, A. V. (2021). Genetic and demographic history define a conservation strategy for earth's most endangered pinniped, the Mediterranean monk seal Monachus monachus. *Scientific reports*, *11*(1), 1-10.

Keller, L. F., Biebach, I., Ewing, S. R., & Hoeck, P. E. (2012). The genetics of reintroductions: inbreeding and genetic drift. *Reintroduction biology: integrating science and management*, 360-394.

Keller, D., Holderegger, R., van Strien, M. J., & Bolliger, J. (2015). How to make landscape genetics beneficial for conservation management?. *Conservation Genetics*, *16*, 503-512.

Kleinman-Ruiz, D., Soriano, L., Casas-Marce, M., Szychta, C., Sánchez, I., Fernández, J., & Godoy, J. A. (2019). Genetic evaluation of the Iberian lynx ex situ conservation programme. *Heredity*, *123*(5), 647-661.

Lacy, R. C., Ballou, J. D., & Pollak, J. P. (2012). PMx: software package for demographic and genetic analysis and management of pedigreed populations. *Methods in Ecology and Evolution*, *3*(2), 433-437.

Lande, R. (1988). Genetics and demography in biological conservation. *Science*, *241*(4872), 1455-1460.

Leite, K. C. E., Seixas, G. H. F., Berkunsky, I., Collevatti, R. G., & Caparroz, R. (2008). Population genetic structure of the blue-fronted Amazon (Amazona aestiva, Psittacidae: Aves) based on nuclear microsatellite loci: implications for conservation. *Genet Mol Res*, *7*, 819-829.

Lewis, T. C., Vargas, I. G., Williams, S., Beckerman, A., & Childs, D. Z. (2022). Using passive acoustic monitoring to estimate the abundance of wide-ranging species in tropical rainforest ecosystems. *bioRxiv*, 2022-12.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in ecology & evolution*, *18*(4), 189-197

Masel, J. (2011). Genetic drift. *Current Biology*, *21*(20), R837-R838

Marshall, T. C., Slate, J. B. K. E., Kruuk, L. E. B., & Pemberton, J. M. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular ecology*, *7*(5), 639-655.

McGreevy Jr, T. J., Dabek, L., & Husband, T. P. (2011). Genetic evaluation of the Association of Zoos and Aquariums Matschie's tree Kangaroo (Dendrolagus matschiei) captive breeding program. *Zoo biology*, *30*(6), 636-646.

Miller, K. A., Chapple, D. G., Towns, D. R., Ritchie, P. A., & Nelson, N. J. (2009). Assessing genetic diversity for conservation management: a case study of a threatened reptile. *Animal Conservation*, *12*(2), 163-171.

Miller-Butterworth, C. M., Vacco, K., Russell, A. L., & Gaspard III, J. C. (2021). Genetic Diversity and Relatedness among Captive African Painted Dogs in North America. *Genes*, *12*(10), 1463.

Mock, K. E., Latch, E. K., & Rhodes, O. E. (2004). Assessing losses of genetic diversity due to translocation: long-term case histories in Merriam's turkey (Meleagris gallopavo merriami). *Conservation Genetics*, *5*(5), 631.

Monge GA, Chassot O, Powell GVN, Palminteri S, Alemán-Zelaya U, Wright P. 2003. Ecologia de la lapa verde (Ara ambiguus) en Costa Rica. Zeledonia. 7(2):4–12.

Moore, J. A., Nelson, N. J., Keall, S. N., & Daugherty, C. H. (2008). Implications of social dominance and multiple paternity for the genetic diversity of a captive-bred reptile population (tuatara). *Conservation Genetics*, *9*, 1243-1251.

Morrison, C. E., Johnson, R. N., Grueber, C. E., & Hogg, C. J. (2020). Genetic impacts of conservation management actions in a critically endangered parrot species. *Conservation Genetics*, *21*, 869-877.

Oklander Tzika, A. C., Remy, C., Gibson, R., & Milinkovitch, M. C. (2009). Molecular genetic analysis of a captive-breeding program: the vulnerable endemic Jamaican yellow boa. *Conservation Genetics*, *10*, 69-77.

Ogden, R., & Linacre, A. (2015). Wildlife forensic science: a review of genetic geographic origin assignment. *Forensic Science International: Genetics*, *18*, 152-159.

Overbeek, A., Galla, S., Brown, L., Cleland, S., Thyne, C., Maloney, R., & Steeves, T. (2020). Pedigree validation using genetic markers in an intensively-managed taonga species, the critically endangered kakī (Himantopus novaezelandiae). *Notornis*.

Pacioni, C., Wayne, A. F., & Spencer, P. B. (2013). Genetic outcomes from the translocations of the critically endangered woylie. *Current Zoology*, *59*(3), 294-310.

Pacioni, C., Trocini, S., Wayne, A. F., Rafferty, C., & Page, M. (2020). Integrating population genetics in an adaptive management framework to inform management strategies. *Biodiversity and Conservation*, *29*(3), 947-966.

Papadopulos, A. S., Kaye, M., Devaux, C., Hipperson, H., Lighten, J., Dunning, L. T., ... & Savolainen, V. (2014). Evaluation of genetic isolation within an island flora reveals unusually widespread local adaptation and supports sympatric speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1648), 20130342.

Pasachnik, S. A., Colosimo, G., Carreras-De Leon, R., & Gerber, G. (2020). Genetic structure of Rhinoceros Rock Iguanas, Cyclura cornuta, in the Dominican Republic, with insights into the impact of captive facilities and the taxonomic status of Cyclura on Mona Island. *Conservation Genetics*, *21*, 837-851.

Seddon, P. J., Griffiths, C. J., Soorae, P. S., & Armstrong, D. P. (2014). Reversing defaunation: restoring species in a changing world. *Science*, *345*(6195), 406-412.

Presti, F. T., & Wasko, A. P. (2014). A review of microsatellite markers and their application on genetic diversity studies in parrots. *Open Journal of Genetics*, 69-77.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.

Ramirez, O., Altet, L., Ensenat, C., Vila, C., Sanchez, A., & Ruiz, A. (2006). Genetic assessment of the Iberian wolf Canis lupus signatus captive breeding program. *Conservation Genetics*, *7*, 861-878

Raymond M, Rousset F (1995) genepop, version 1.2: population genetics software for exact tests and ecumenicism. Journal of Heredity, 86, 248–249.

Redford, K. H., Jensen, D. B., & Breheny, J. J. (2012). Integrating the captive and the wild. *Science*, *338*(6111), 1157-1158

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular ecology resources*, *8*(1), 103-106.

Russello, M. A., Stahala, C., Lalonde, D., Schmidt, K. L., & Amato, G. (2010). Cryptic diversity and conservation units in the Bahama parrot. *Conservation Genetics*, *11*, 1809-1821.

Russello, M. A., & Amato, G. (2004). Ex situ population management in the absence of pedigree information. *Molecular ecology*, *13*(9), 2829-2840.

Sánchez-Ramírez, S., Rico, Y., Berry, K. H., Edwards, T., Karl, A. E., Henen, B. T., & Murphy, R. W. (2018). Landscape limits gene flow and drives population structure in Agassiz's desert tortoise (Gopherus agassizii). *Scientific Reports*, *8*(1), 11231

Seutin, G., White, B. N., & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian journal of zoology*, *69*(1), 82-90

Schreier, A., Stephenson, S., Rust, P., & Young, S. (2015). The case of the endangered Kootenai River white sturgeon (Acipenser transmontanus) highlights the importance of post-release genetic monitoring in captive and supportive breeding programs. *Biological Conservation*, *192*, 74-81.

Stojanovic, D., Olah, G., Webb, M., Peakall, R., & Heinsohn, R. (2018). Genetic evidence confirms severe extinction risk for critically endangered swift parrots: implications for conservation management. *Animal Conservation*, *21*(4), 313-323.

Tallmon, D. A., Luikart, G., & Waples, R. S. (2004). The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution*, *19*(9), 489-496.

Taylor, H. R., Robertson, H., Carter, A. L., & Ramstad, K. M. (2021). The conservation management implications of isolation by distance and high genetic diversity in Great Spotted Kiwi (Apteryx haastii). *Emu-Austral Ornithology*, *121*(1-2), 10-22.

Tiffin, P., & Ross-Ibarra, J. (2014). Advances and limits of using population genetics to understand local adaptation. *Trends in ecology & evolution*, *29*(12), 673-680.

Townsend, S. M., & Jamieson, I. G. (2013). Molecular and pedigree measures of relatedness provide similar estimates of inbreeding depression in a bottlenecked population. *Journal of Evolutionary Biology*, *26*(4), 889-899.

Verma, A. K. (2016). Biodiversity: Its different levels and values. *International Journal on Environmental Sciences*, *7*(2), 143-145.

Wang, J. (2011). COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular ecology resources*, *11*(1), 141-145.

Weeks AR, Sgro CM, Young AG, Frankham R, Mitchell NJ, Miller KA, Byrne M, Coates DJ, Eldridge MDB, Sunnucks P, Breed MF, James EA, Hoffmann AA (2011) Assessing the benefits and risks of translocations in changing environments: a genetic perspective. Evol Appl 4(6):709–725.

Weng, Z., Yang, Y., Wang, X., Wu, L., Hua, S., Zhang, H., & Meng, Z. (2021). Parentage analysis in giant grouper (Epinephelus lanceolatus) using microsatellite and SNP markers from genotyping-by-sequencing data. *Genes*, *12*(7), 1042.

West, R., Potter, S., Taggart, D., & Eldridge, M. D. (2018). Looking back to go forward: genetics informs future management of captive and reintroduced populations of the black-footed rock-wallaby Petrogale lateralis. *Conservation Genetics*, *19*, 235-247.

White, D. J., Watts, C., Allwood, J., Prada, D., Stringer, I., Thornburrow, D., & Buckley, T. R. (2017). Population history and genetic bottlenecks in translocated Cook Strait giant weta, Deinacrida rugosa: recommendations for future conservation management. *Conservation genetics*, *18*, 411-422.

White, N. F., Mennell, H., Power, G., Edwards, D., Chrimes, L., Woolaver, L., ... & Orozco-terWengel, P. (2022). A population genetic analysis of the Critically Endangered Madagascar big-headed turtle, Erymnochelys madagascariensis across captive and wild populations. *Scientific Reports*, *12*(1), 8740.

Whiteley, A. R., Fitzpatrick, S. W., Funk, W. C., & Tallmon, D. A. (2015). Genetic rescue to the rescue. *Trends in ecology & evolution*, *30*(1), 42-49.

Willi, Y., Kristensen, T. N., Sgrò, C. M., Weeks, A. R., Ørsted, M., & Hoffmann, A. A. (2022). Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. *Proceedings of the National Academy of Sciences*, *119*(1), e2105076119.

Willoughby, J. R., Fernandez, N. B., Lamb, M. C., Ivy, J. A., Lacy, R. C., & DeWoody, J. A. (2015). The impacts of inbreeding, drift and selection on genetic diversity in captive breeding populations. Molecular Ecology, 24(1), 98–110. https://doi.org/10.1111/mec.13020

Witzenberger, K. A., & Hochkirch, A. (2011). Ex situ conservation genetics: a review of molecular studies on the genetic consequences of captive breeding programmes for endangered animal species. *Biodiversity and conservation*, *20*, 1843-1861.

Witzenberger, K. A., & Hochkirch, A. (2013). Evaluating ex situ conservation projects: Genetic structure of the captive population of the Arabian sand cat. *Mammalian Biology*, *78*, 379-382

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

TABLES AND FIGURES

Table 1. 24 microsatellite loci assembled in seven multiplexes and tested on 24 unrelated individuals of the Great Green Macaw from the MRN captive population. M, Multiplex; Min, Minimum allele size; Max, Maximum allele size; Ho, observed heterozygosity; He, expected heterozygosity; F(null), Null alleles; HWE, (p-val) for the HWE test. *** Primers not tested

M	Name	Forward Sequence	Reverse Sequence	Min	Max	Ho	He	F(Null)	HWE
1	Aamb_25191	CTCCAACAGTTTGCAGGTTC	GGGAGTAAACAGCACAGTGG	164	206	0.708	0.615	-0.0811	0.3792
	Aamb_3648	GAGATTGATGCTGTGTATGTCG	ATCTGCCCAGTAGCACTCAG	173	203	0.417	0.355	-0.1086	0.0316
	Aamb_12149	TGTTCAGGTCCAAACCAATC	GATCCCTTCTGCTATCCTGTATC	151	159	0.75	0.784	0.0195	1
2	Aamb_13758	TTTCTCATTGCCTGGAAACC	TCGTAAGAAATATGCTTGGAAGG	137	176	0.5	0.592	0.099	0.8645
	Aamb_24991	CTGCAATGGCACCCTGAC	TCGAGGTTGAATCCAGAGGTC	190	214	0.583	0.487	-0.1087	0.1498
	Aamb_11776	CCATAATGCACATGCTGCTC	TGCAGGAGTTGTAGGAATTGG	**					
3	Aamb_25928	TTCGGTTCCTAGCAAAGAGG	GGGTCAGGCACTGTCTCAG	132	163	0.625	0.707	0.0866	0.0106
	Aamb_2232	CAGCTGAGAAACCTGGAGGAC	CCTGCAACACCTGCAACAC	2*					
	Aamb_2689	AGGTAGCACCAACACTCAGC	GCATAGGTGAGCAGAAGAGG	**					
	Aamb_9690	ACAATTCCCTGCCTGCTC	AGGAAAGTGCTAGAATTGAGACG	173	197	0.583	0.63	0.0216	1
4	Aamb_9238	GCTGTGTTACGCATCTGTGG	AGAAGGTGACCCTGTTGCAC	212	236	0.458	0.467	-0.0009	0.2482
	Aamb_15058	TCAGCATGCCCATGAAATAC	TTTCTTGTGCAGAAACTTCCAC	151	167	0.292	0.361	0.0956	0.555
	Aamb_21942	GATAGACAGGAGGCGGTTTG	AACCAAGTGCTCATTCACCTG	177	185	0.792	0.755	-0.0383	1
5	Aamb_18911	GAGCCAGATTTATGAGCATTTG	GCCATGAGCTCAAGAGACAG	191	236	0.458	0.731	0.2179	0.579
	Aamb_15017	GTGCATGCCTTGACTTGTG	TGCATATTGCAATGAAGTATATGG	221	231	0.583	0.507	-0.0804	0.6775
	Aamb_3822	TCCATGATTGTATGGGAGTTTG	AGAAGTTTCAGGGCCATCTG	195	245	0.458	0.731	0.2179	0.0078
6	Aamb_4366	TCCGTGTTTGAAGGTGAACTC	ACCAACATTAGGCTGGATGC	218	242	0.583	0.469	-0.153	0.6111

	Aamb_5826	CATCATCTGTGAGGCAGCAG	TGTTGAGCTCTAGACAGCATTCC	241	247	0.625	0.621	-0.0216	1
	Aamb_7958	CATGTCCTGGCACCAACC	CTTTCCGTCTGCATTTCCTG	183	229	0.125	0.12	-0.0234	0.7234
	Aamb_24939	AGGACACCTGACCCAAACTG	CTCACCGCCTAATACCAAGC	**					
7	Aamb_cons_gr3_2	CTAGAGCTAGGAACTGAACACACG	GCTGAGGAGGTTGGACTGAG	**					
	Aamb_13122	AGCTTGGAATCCTCAGCTTG	AGCTAGGGAAGTGTCGCATC	**					
	Aamb_25306	TCCACTTCCTCATCCAAAGG	ATGGTGGGTGTCAGGTGTG	**					

1158

Table 2. Mean and standard deviation of genetic diversity per population; N, sample size; Ho, observed heterozygosity; He, expected heterozygosity; Ar, allelic richness, of all the sampled populations.

Population	***N***	***Ar***	***Private alleles***	***Ho(sd)***	***He (sd)***	***TrioML (sd)***
Wild	13	3.62	0	0.548 (0.26)	0.530 (0.19)	0.1482 (0.252)
MRN	69	5.5	11	0.552 (0.20)	0.537 (0.19)	0.0825 (0.109)
Zooaves	22	3.62	2	0.558 (0.24)	0.517 (0.18)	0.1086 (0.132)
Colombia	13	3.75	11	0.767 (0.27)	0.612 (0.18)	0.1206 (0.086)
Europe	22	4.75	5	0.680 (0.16)	0.622 (0.16)	0.0670 (0.116)
GLOBAL	138		-	0.613 (0.17)	0.559 (0.15)	0.0941 (0.133)

Table 3. Estimated relatedness of previously unknown relation between breeding pairs using the likelihood estimator TrioML

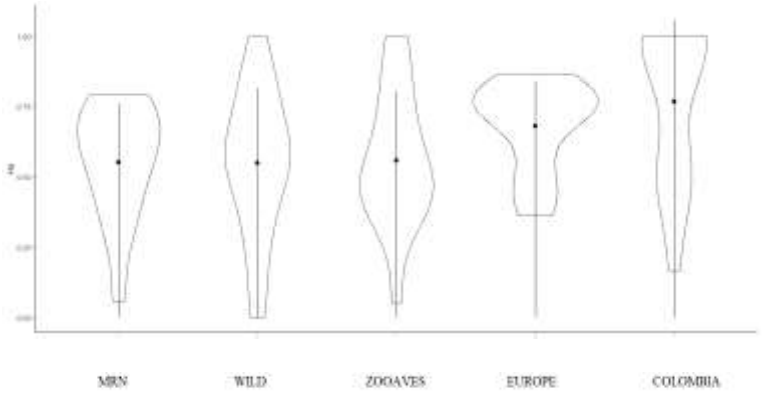
ID	Breeding Pair	Estimated Relatedness (R)
1	ARA41 + RM123	0.0
2	RM136 + RICH6	0.264
3	RICH9 + RM129	0.0418
4	RM125 + RICH119	0.0
5	RM319 + RM339	0.05
6	ARA20 + RM137	0.05
7	ARA22 + ARA23	0.0
8	RM325 + RM341	0.025
9	RM312 + RICH15	0.18

Table 4. Mean and standard deviation of genetic diversity per population; N, sample size; Ho, observed heterozygosity; He, expected heterozygosity; Ar, allelic richness and Relatedness estimate for the Costa Rica release cohort and captive populations.

Population	N	Ar	Ho(sd)	He (sd)	Relatedness Estimate
Release	27	3.27	0.5531 (0.229)	0.5248 (0.198)	0.092
CR Captive	62	4.25	0.5454 (0.182)	0.5318 (0.195)	0.11

Figure 1. Distribution map for the Great Green Macaw (Ara ambiguus), ranging from Honduras to Ecuador. The colored dots represent locations where samples were collected for the captive and wild populations, including an insert of the European captive population (EEP).

(a)



(b)

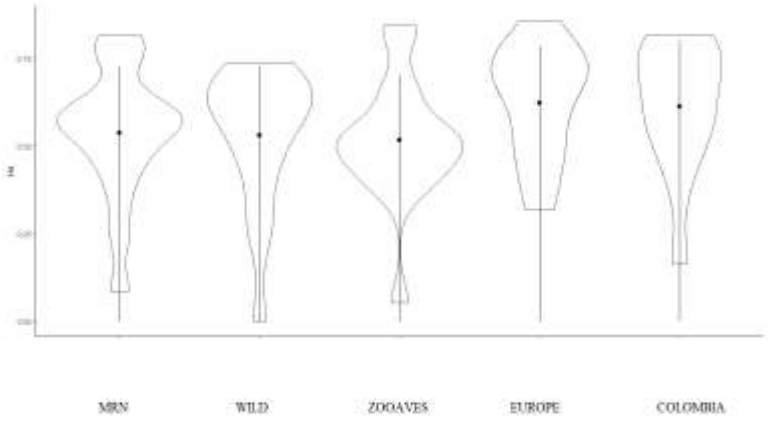
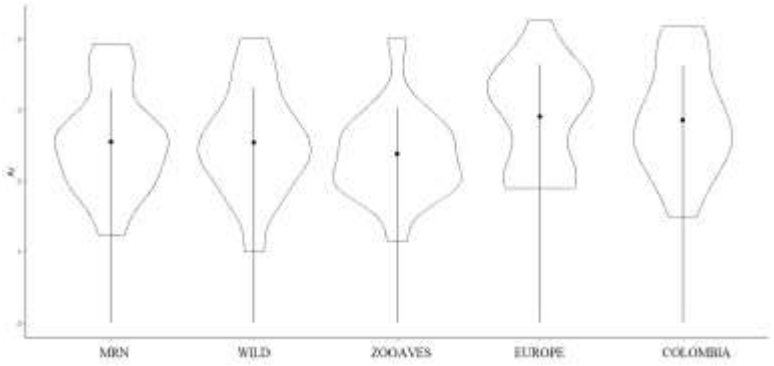


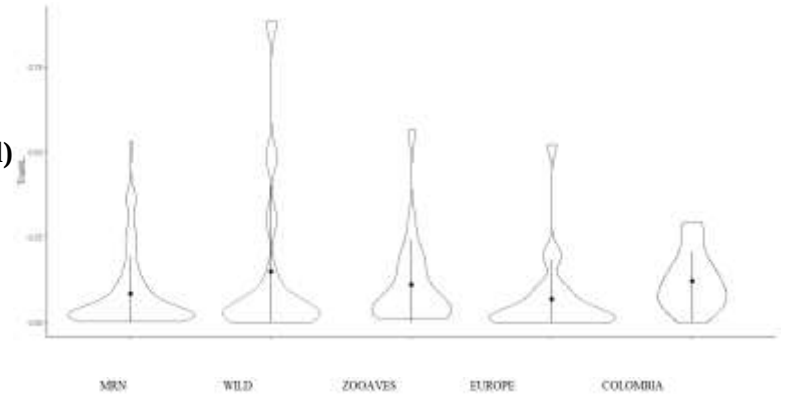
Figure 2. Distribution of genetic diversity and inbreeding data, and mean value per population n; 2.a Ho, observed heterozygosity; 2.b He, expected heterozygosity; 2.c Ar, allelic richness, 2.d Trio, inbreeding coefficient (sample sizes: wild, n=13; MRN, n=69; ZA, n=20; Colombia, n=13; European, n=22).

(c)



1191

(d)

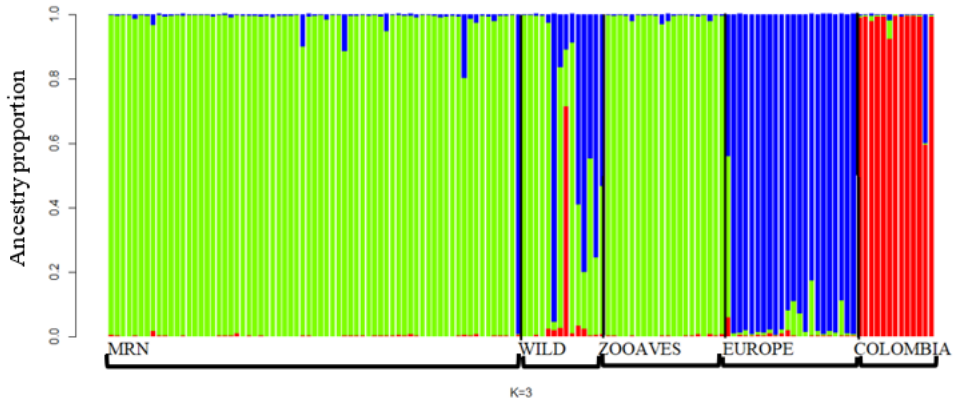


1192

1193

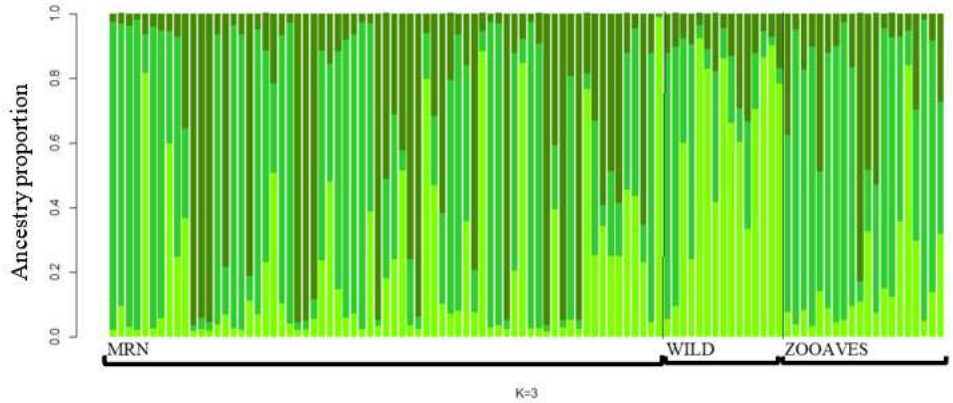
1194

(a)


Ancestry proportion
0.0
0.2
0.4
0.6
0.8
1.0
MRN
WILD
ZOOAVES
EUROPE
COLOMBIA
K=3

1195

(b)


Ancestry proportion
0.0
0.2
0.4
0.6
0.8
1.0
MRN
WILD
ZOOAVES
K=3

1196

(c)

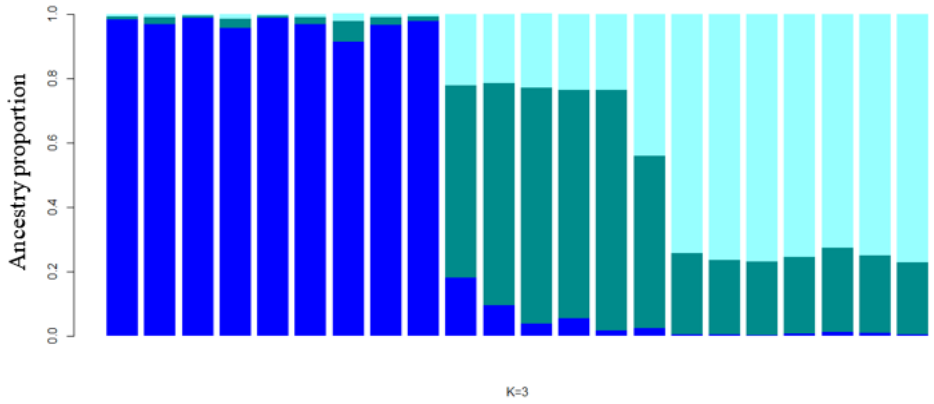


Figure 3. STRUCTURE outputs of based on analyses of different subsets of the sampled Great Green Macaw populations. (**a**) K=5 for all sampled populations, where Cluster 1 is represented in red, cluster 2 is represented in green and cluster 3 is represented in blue. (**b**) K=3 for Costa Rican samples, populations of origin of the samples named between brackets (population of origin of the samples named between brackets), where cluster 1 is represented by lime green, cluster 2 by dark green and cluster 3 by moss green, and (**c**) K=3 for European sample

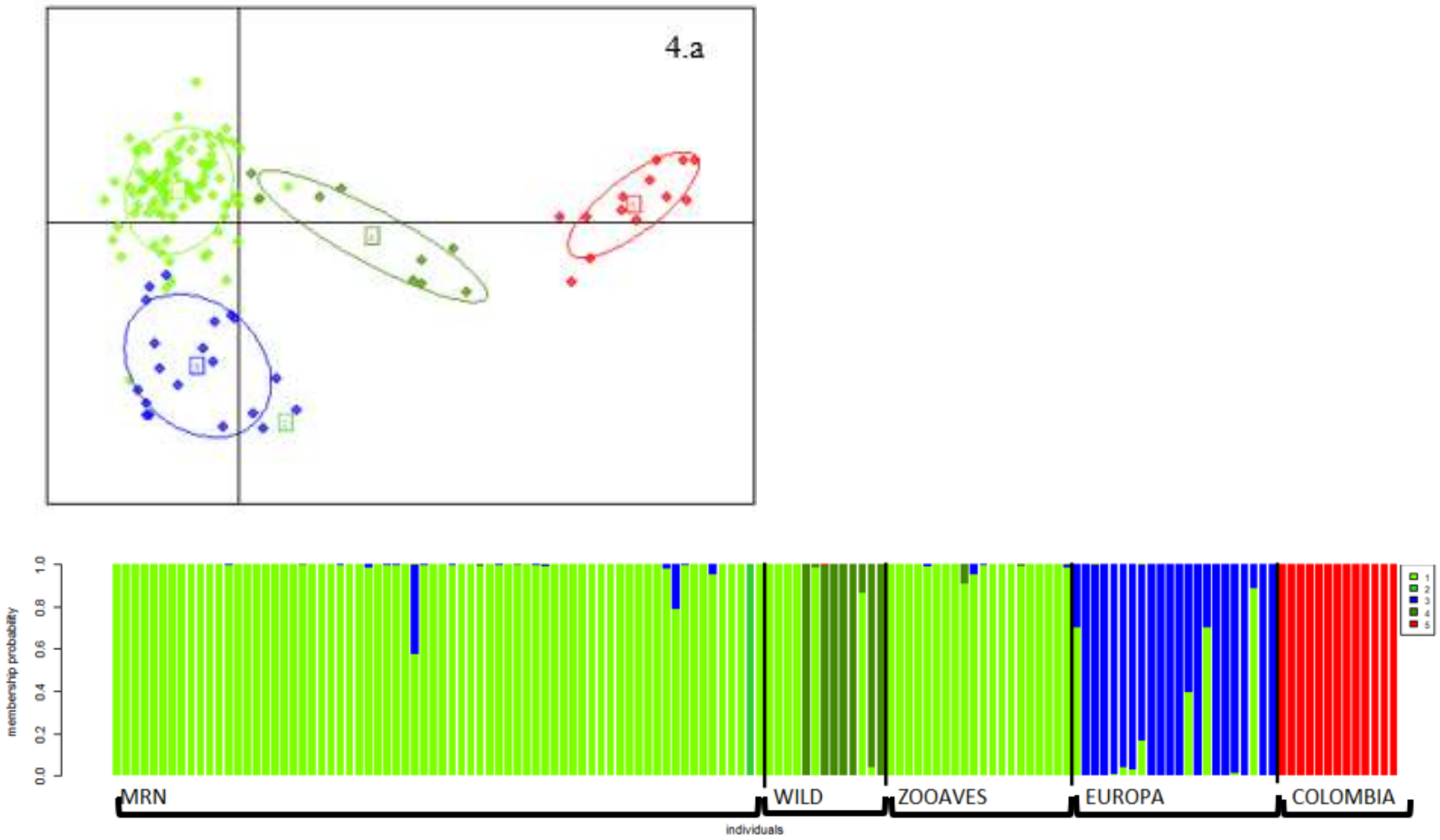


Figure 4. DAPC multivariate analysis of Great Green Macaw genotype data, populations of origin of the samples named between brackets. DAPC K = 5 scatterplot of all individual macaws assigned to inferred 5 subpopulation clusters considering (a) PC1 and PC2, (b) DAPC barplot representing probability of assignment of individuals to each group.

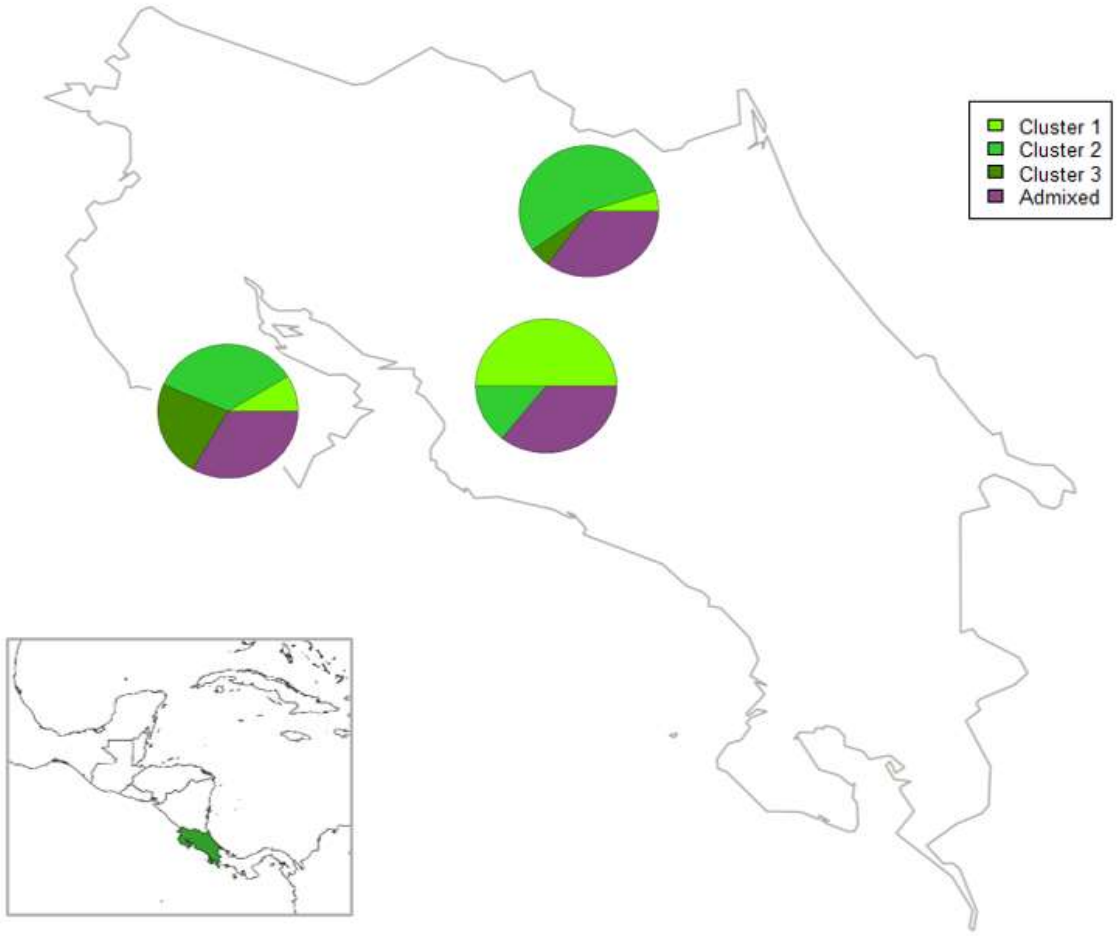


Figure 5 Map with the percentages of each Costa Rican population per cluster. Insert represents Central America, with Costa Rica highlighted in green. The populations from right to left correspond to Macaw Recovery Network, Wild and Zooaves.

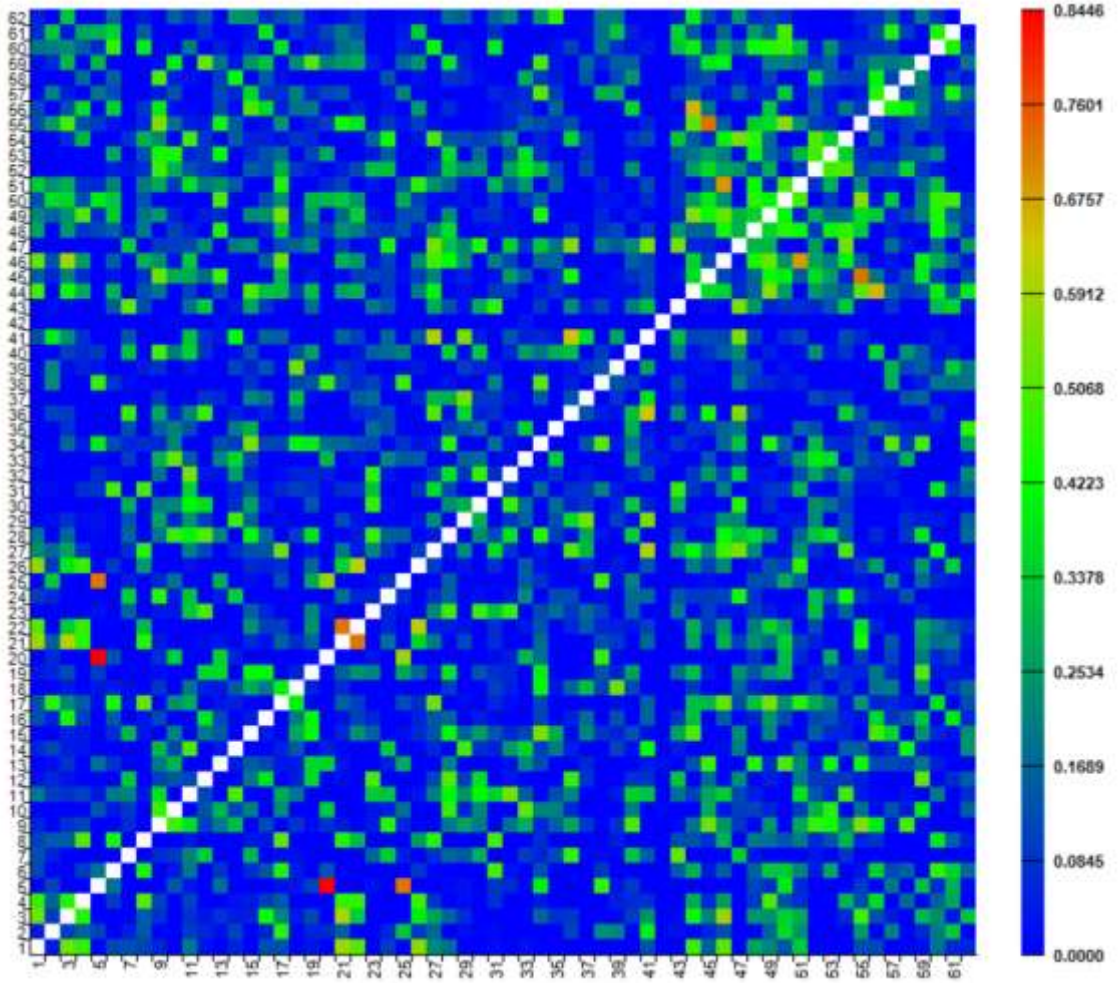


Figure 6. Relatedness Plot. All captive individuals from MRN and Zooaves that are not going to be released. The pairwise estimated relatedness values are represented in the matrix as colors from the scale, which ranges from blue (r=0) to red (r=0.8846).

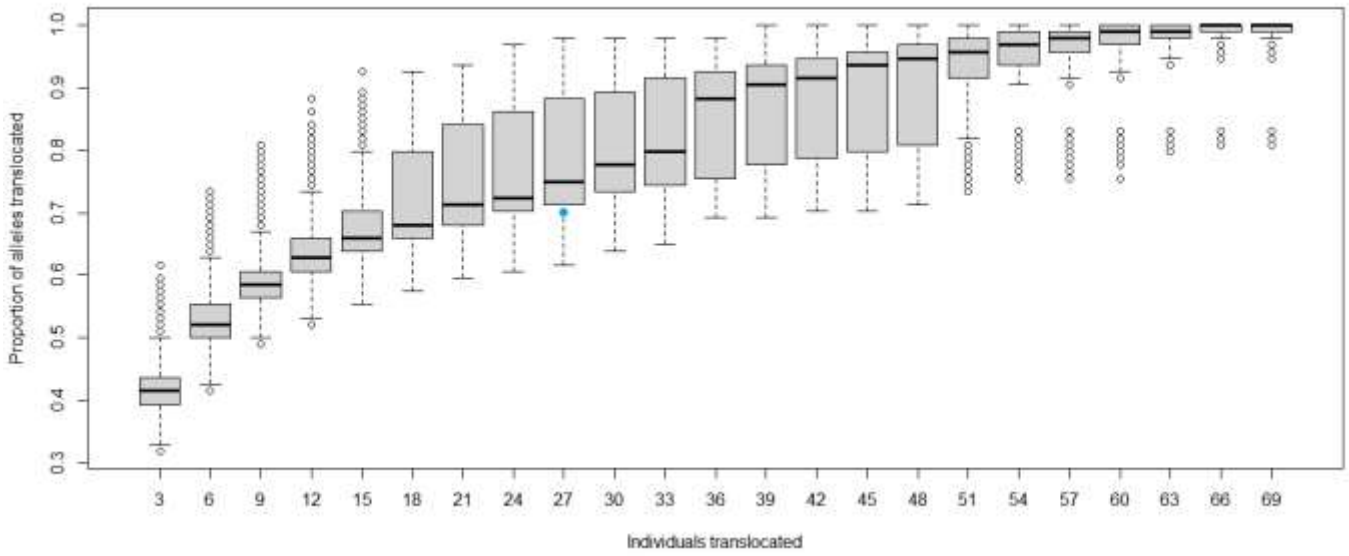


Figure 7. Proportion of alleles in the MRN source population of Great Green Macaws represented by in the translocation of different numbers of individuals. Black point represents the values obtained in the 100 replicate runs for each number of individuals transferred, with the mean of each number of simulated individuals represented in blue. The blue dot indicates the proportion of alleles that are present in the selected MRN release population

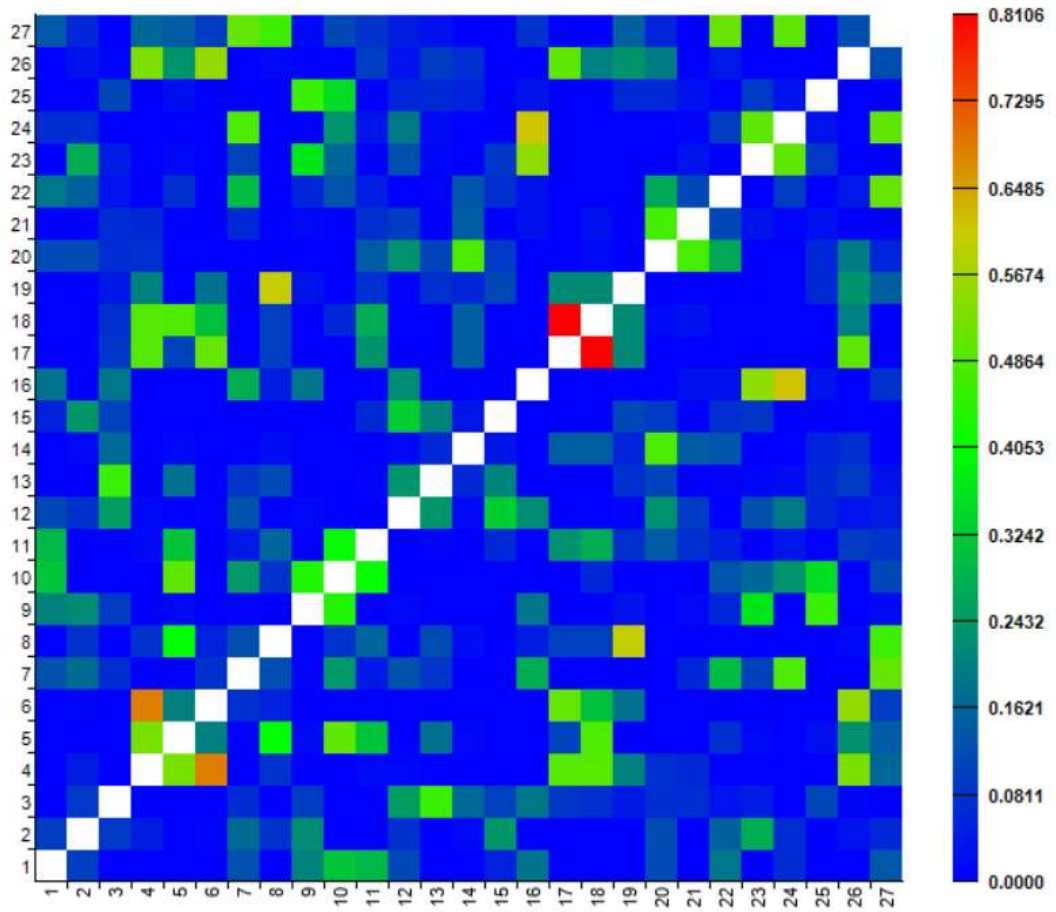
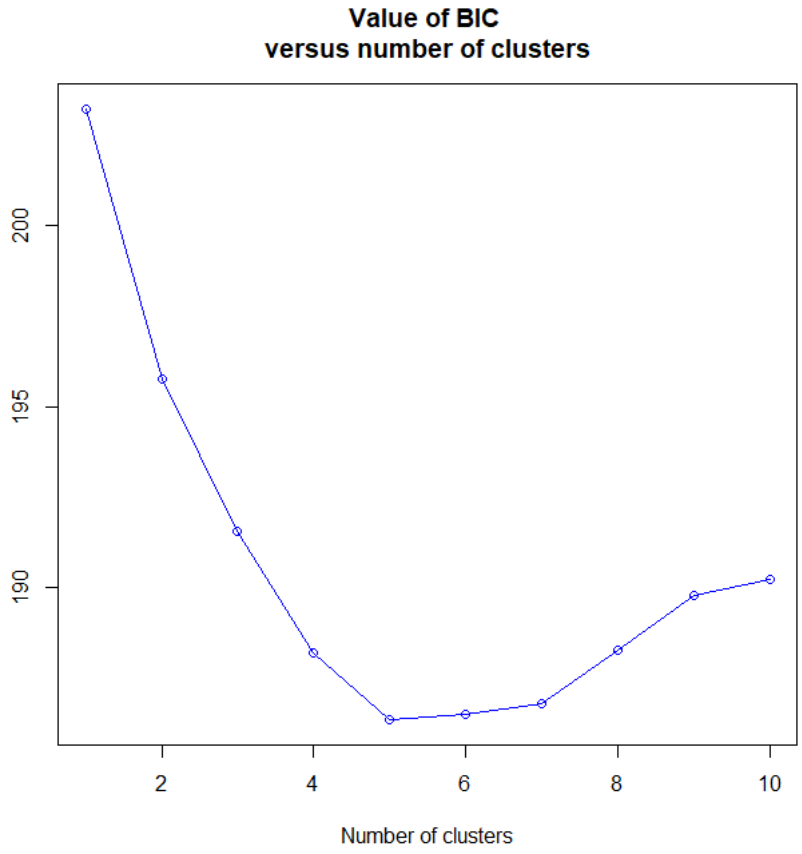


Figure 8. Relatedness plot. Relatedness between MRN individuals that are being considered for release as a founding population. The pairwise estimated relatedness values are represented in the matrix as colors from the scale, which ranges from blue (r=0) to red (r=0.8846)

ANEX I. Suplementary information

(a)



(b)

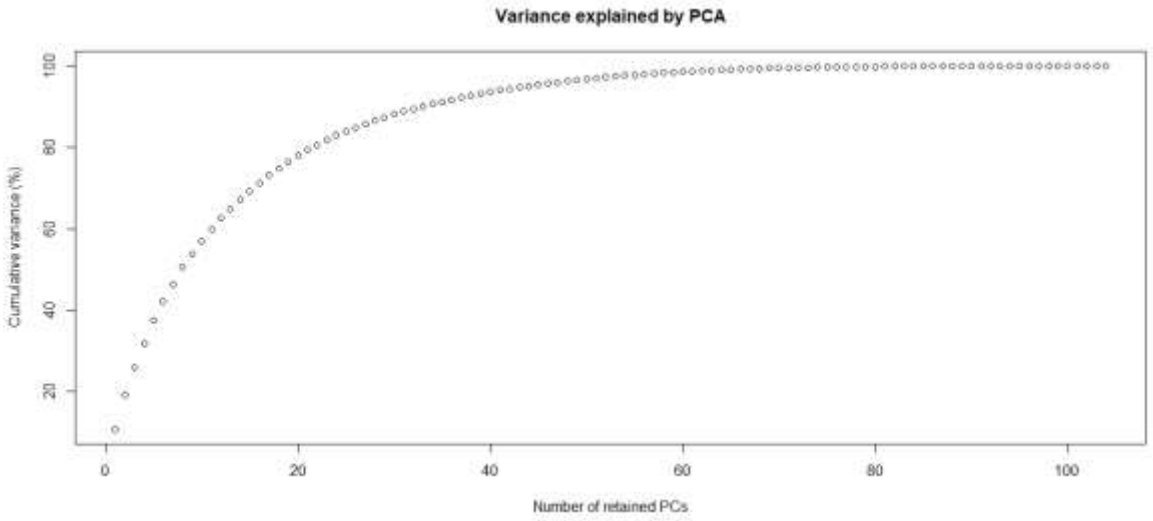
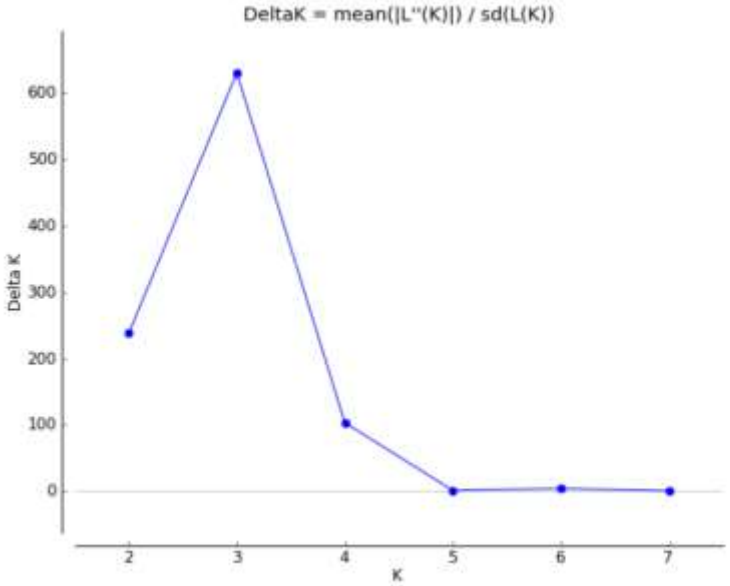


Figure S-1. Graphic output of the *find.cluster* function in the Adegent package; 1.a Bayesian Information Criterion approach shows the break in the number of clusters, in this case in K=5; 1.b Variance accumulated by the number of PCA retained.

(a)

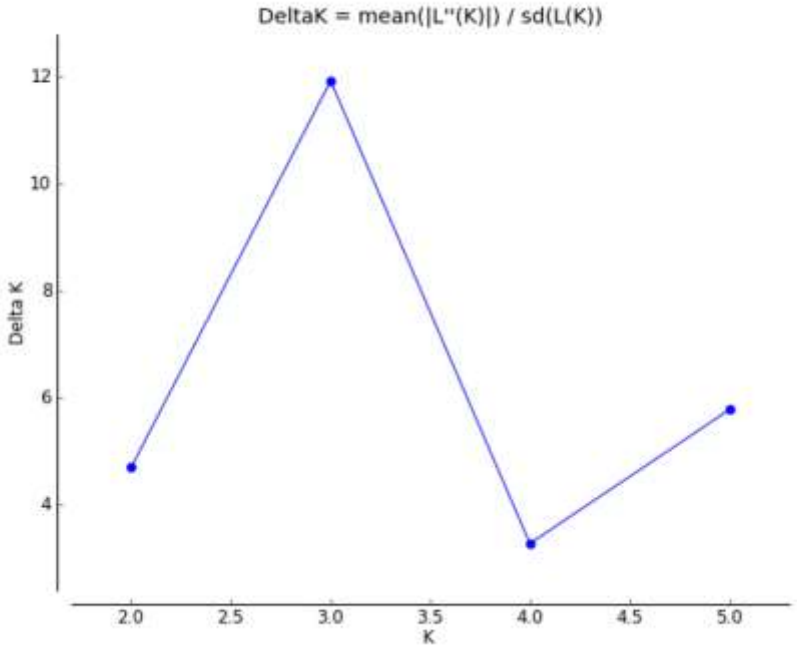


(b)

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	7	-4886.542857	0.181265	—	—	—
2	7	-4448.214286	0.971989	438.328571	231.042857	237.701192
3	7	-4240.928571	0.213809	207.285714	134.757143	630.268824
4	7	-4168.400000	2.257580	72.528571	232.657143	103.055998
5	7	-4328.528571	361.975398	-160.128571	391.800000	1.082394
6	7	-4096.857143	50.570342	231.671429	199.100000	3.937090
7	7	-4064.285714	7.649058	32.571429	4.357143	0.569631
8	7	-4036.071429	5.359638	28.214286	—	—

Figure S2. Result of the STRUCTURE HARVESTER analysis for the 5 populations, with a setup of 7 with independent simulations, 1,000,000 iterations with a burn-in of 200,000 for K=8; 2.a Graphic representation of Delta of K (as depicted in the formula) in function of different numbers of clusters, the chosen K corresponds to the highest Delta of K; 2.b Table resulting from the analysis showing the highest Delta of K in K=3.

(a)

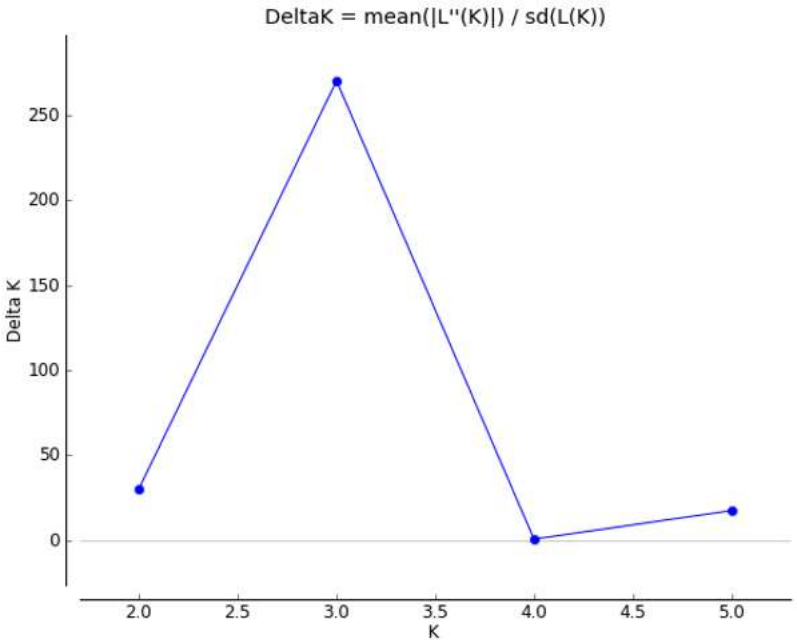


(b)

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	7	-2976.928571	0.540723	—	—	—
2	7	-2909.971429	3.633508	66.957143	17.042857	4.690469
3	7	-2860.057143	2.194582	49.914286	26.171429	11.925473
4	7	-2836.314286	2.056233	23.742857	6.714286	3.265333
5	7	-2819.285714	5.216457	17.028571	30.100000	5.770199
6	7	-2832.357143	7.376281	-13.071429	—	—

Figure S3. Result of the STRUCTURE HARVESTER analysis for the Costa Rican populations, with a setup of 7 with independent simulations, 1,000,000 iterations with a burn-in of 200,000 for K=6; 2.a Graphic representation of Delta of K (as depicted in the formula) in function of different numbers of clusters, the chosen K corresponds to the highest Delta of K; 2.b Table resulting from the analysis showing the highest Delta of K in K=3.

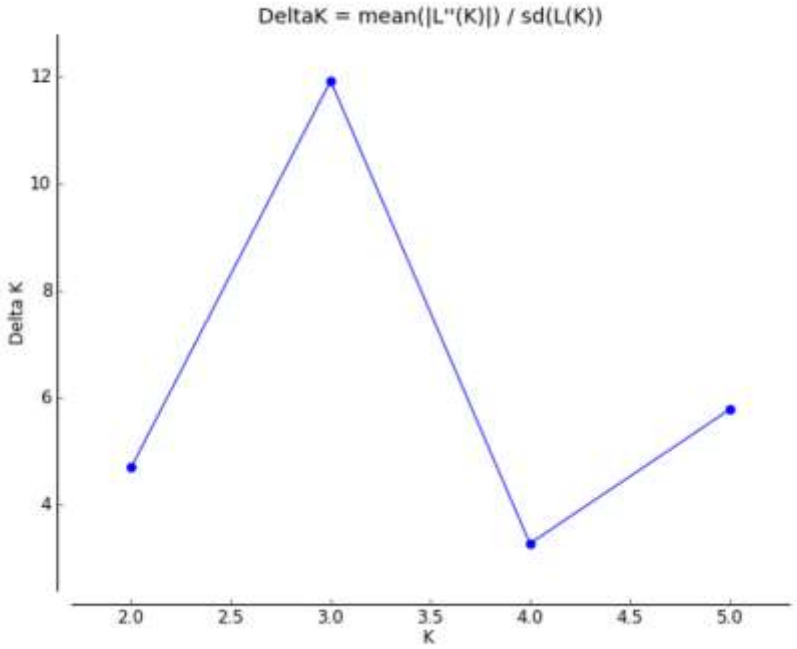
(a)



(b)

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	7	-847.600000	0.939858	—	—	—
2	7	-792.085714	1.393352	55.514286	42.642857	30.604521
3	7	-779.214286	0.823465	12.871429	222.314286	269.974051
4	7	-988.657143	488.146842	-209.442857	367.028571	0.751881
5	7	-831.071429	10.223293	157.585714	178.771429	17.486678
6	7	-852.257143	9.237398	-21.185714	—	—

Figure S4. Result of the STRUCTURE HARVESTER analysis for the European samples, with a setup of 7 with independent simulations, 1,000,000 iterations with a burn-in of 200,000 for K=6; 2.a Graphic representation of Delta of K (as depicted in the formula) in function of different numbers of clusters, the chosen K corresponds to the highest Delta of K; 2.b Table resulting from the analysis showing the highest Delta of K in K=3.



K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	7	-2976.928571	0.540723	—	—	—
2	7	-2909.971429	3.633508	66.957143	17.042857	4.690469
3	7	-2860.057143	2.194582	49.914286	26.171429	11.925473
4	7	-2836.314286	2.056233	23.742857	6.714286	3.265333
5	7	-2819.285714	5.216457	17.028571	30.100000	5.770199
6	7	-2832.357143	7.376281	-13.071429	—	—

Figure S3. Result of the STRUCTURE HARVESTER analysis with a setup of 7 with independent simulations, 1,000,000 iterations with a burn-in of 200,000 for K=6; 2.a Graphic representation of Delta of K (as depicted in the formula) in function of different numbers of clusters, the chosen K corresponds to the highest Delta of K; 2.b Table resulting from the analysis showing the highest Delta of K in K=3.

Relatedness Coefficient	Correlation
TrioML	0.764
Wang	0.730
LynchLi	0.729
LinchRd	0.712
Ritland	0.237
Queller GT	0.744
DyadML	

Table S1. Estimated Relatedness correlation coefficient to “true” estimated relatedness value in the simulation performed in COANCESTRY.