1   **Identification of Genetically Important Individuals of the Rediscovered Floreana**

2   **Galápagos Giant Tortoise (*Chelonoidis elephantopus*) Provide Founders for Species**

3   **Restoration Program**

4   Joshua M. Miller[1†]*, Maud C. Quinzin[1†], Nikos Poulakakis[2,3], James P. Gibbs[4], Luciano B.

5   Beheregaray[5], Ryan C. Garrick[6], Michael A. Russello[7], Claudio Ciofi[8], Danielle L. Edwards[9],

6   Elizabeth A. Hunter[10], Washington Tapia[11,12], Danny Rueda[12], Jorge Carrión[12], Andrés A.

7   Valdivieso[1], Adalgisa Caccone[1]*

8

9   1. Department of Ecology and Evolutionary Biology, Yale University, 21 Sachem St. New

10   Haven, Connecticut, 06520, United States of America

11   2. Department of Biology, School of Sciences and Engineering, University of Crete,

12   Vasilika Vouton, Gr-71300, Heraklio, Crete, Greece

13   3. Natural History Museum of Crete, School of Sciences and Engineering, University of

14   Crete, Knossos Av., GR-71409, Heraklio, Crete, Greece

15   4. College of Environmental Science & Forestry, State University of New York, Syracuse,

16   New York, 13210, United States of America

17   5. Molecular Ecology Lab, School of Biological Sciences, Flinders University, GPO

18   Box 2100, Adelaide, SA, 5001, Australia

19   6. Department of Biology, University of Mississippi, Oxford, Mississippi, 38677, United

20   States of America

21   7. Department of Biology, University of British Columbia, Okanagan Campus, Kelowna,

22   BC V1V 1V7, Canada

23   8. Department Biology, University of Florence, 50019 Sesto Fiorentino (FI), Italy

24   9. Life and Environmental Sciences, University of California, Merced, 5200 N Lake Rd,

25   Merced, California, 95343, United States of America

26   10. Department of Natural Resources and Environmental Science, University of Nevada –

27   Reno, Max Fleischmann Agricultural Building, Reno NV 89557 USA

28   11. Galapagos Conservancy, Fairfax, Virginia, 22030, United States of America

29   12. Galápagos National Park Directorate, Puerto Ayora, Galápagos, Ecuador

30   *Corresponding authors: joshua.miller@yale.edu, adalgisa.caccone@yale.edu

31   †These authors contributed equally to this work

32 **Abstract**

33 Species are being lost at an unprecedented rate due to human-driven environmental changes. The

34 cases in which species declared extinct can be revived are rare. However, here we report that a

35 remote volcano in the Galápagos Islands hosts many giant tortoises with high ancestry from a

36 species previously declared as extinct: *Chelonoidis elephantopus* or the Floreana tortoise. Of 150

37 individuals with distinctive morphology sampled from the volcano, genetic analyses revealed

38 that 65 had *C. elephantopus* ancestry and thirty-two were translocated from the volcano's slopes

39 to a captive breeding center. A genetically informed captive breeding program now being

40 initiated will, over the next decades, return *C. elephantopu*s tortoises to Floreana Island to serve

41 as engineers of the island's ecosystems. Ironically, it was the haphazard translocations by

42 mariners killing tortoises for food centuries ago that created the unique opportunity to revive this

43 "lost" species today.

44

45 **Introduction**

46 Human activities have generated an extreme and rapid loss of biodiversity[1]. Many actions

47 have been undertaken to prevent species extinctions, including creating laws to protect

48 endangered species and critical habitats[2], translocations of individuals among populations or into

49 new habitats[3,4], and *ex situ* management including captive breeding[5]. For many species, such

50 actions come too late to facilitate recovery[6–8]. Although 'de-extinction' using laboratory

51 techniques is currently being debated[9], such methods are viable only for some taxa and will

52 generate significant anticipated, and unanticipated, risks[10,11]. Generally, extinction is final and

53 cases where lost species can be revived will be extremely rare.

54 Despite their insularity, even remote oceanic islands are not exempt from rapid

55 anthropogenic changes. For example, the ecosystems of the Galápagos Islands, located ~900

56 kilometers off the Pacific coast of Ecuador, have been degraded by human activities since the

57 archipelago was discovered in 1535. The resulting loss of biodiversity, including many endemic

58 species, has been chiefly due to the introduction of non-native species[12,13]. In response, concerted

59 efforts have been taken to restore ecosystems on the islands, including removal of introduced

60 pests[14,15] and population restoration via captive breeding and repatriation of threatened native

61 species (e.g. [16,17]).

2

62   Galápagos giant tortoises (*Chelonoidis* spp.) are flagship species for ongoing restoration
63 efforts[18] in this archipelago and play an important functional role as mega-herbivores in the
64 islands' ecosystems[19]. Galápagos giant tortoises can be classified into 15 species based on
65 genetic data[20]. Generally, there is a single species per island with two exceptions: Isabela Island
66 that has a different endemic species associated with each of its five volcanoes, and Santa Cruz
67 Island that contains two species, one recently described[20] (Figure 1a). The 15 species exhibit two
68 general carapace shapes: "domed," a rounded cupola-like form (Figure 1c), and a "saddle-
69 backed" form, with a high anterior opening creating the shape of a saddle (Figure 1d). Five of the
70 15 species have the saddle-backed morphology: *C. elephantopus* (Floreana), *C. hoodensis*
71 (Española), *C. abingdoni* (Pinta), *C. ephippium* (Pinzón), and *C. chathamensis* (San Cristóbal;
72 Figure 1a).

73   Over the past three centuries, all giant tortoise populations experienced a ~90% decline[21],
74 having been killed mostly for food and oil by whalers, sealers, buccaneers, and early
75 colonists[22,23]. Four species have been declared Extinct[24], including two of the five saddle-backed
76 species: *C. elephantopus* from Floreana Island and, most recently, *C. abingdoni,* from Pinta
77 Island. The latter species was represented by a single individual, Lonesome George, until his
78 death in 2012.

79   Surprisingly, recent research found living, wild tortoises with genetic ancestry from two
80 of the extinct saddle-backed species, *C. elephantopus* and *C. abingdoni* (hereafter referred to as
81 the Floreana and Pinta tortoises, respectively) outside their native range[25–28]. These individuals,
82 likely the descendants of tortoises translocated among islands by mariners[22,23,29], were discovered
83 on the remote Volcano Wolf on Isabela Island (Figure 1a,b). Among >1600 tortoises sampled
84 from Volcano Wolf in 2008 during an exploratory expedition, 105 individuals were admixed
85 between the locally endemic species, *C. becki*, and the Floreana ($n = 86$)[5] or Pinta ($n = 17$)[4]
86 species. The majority of these genetically admixed individuals were found on Volcano Wolf's
87 western slopes, facing Banks Bay (also known as Puerto Bravo, PBR: Figure 1b), with a smaller
88 number located on the volcano's northwestern slopes near Piedras Blancas (PBL)[25,26]. No
89 purebred individuals of either of the two non-native species were found in 2008, but genetic
90 simulations and the young age of some mixed ancestry individuals indicated that purebred
91 Floreana and Pinta tortoises might still be present on Volcano Wolf[25,26].

92   Here we build on this previous work. In November 2015 we mounted a 10-day-long

93   search involving ~70 field personnel combined with helicopter- and ship-support. Focusing our

94   search, we explored zones of Volcano Wolf most likely to contain individuals with ancestry from

95   these two "lost" species (Figure 1b) and restricted genetic sampling to those individuals with

96   saddle-backed morphology among the thousands of locally endemic domed *C. becki*. We then

97   assigned ancestry to these tortoises using reference databases containing both extant and extinct

98   species that have previously been used to assign ancestry to tortoises in the wild and captivity[25–

99   31]. Based on these assignments, we determined suitability of the tortoises for a genetically

100   informed captive breeding program aimed at reintroducing these key ecosystem engineers to

101   their native island.

102

103   **Results & Discussion**

104   In total, we encountered 144 individuals with saddle-backed morphology. Of those, 112

105   were released after taking blood samples, and 32 with pronounced saddle-backed morphology

106   were transported to the Galápagos National Park Service's captive tortoise breeding facility on

107   Santa Cruz Island[32–34]. We assigned ancestry to all 144 of these individuals along with six

108   saddle-backed tortoises known to have Floreana ancestry[35] already residing at the breeding

109   center using information from ~700-bp of mitochondrial DNA sequence and diploid genotypes

110   from 12 nuclear microsatellite loci. These loci have previously been shown to accurately assign

111   individuals to tortoise species[25–31].

112   Thirty-five of the 150 individuals analyzed had a mitochondrial DNA haplotype

113   diagnostic of the Floreana species (Supplementary Figure S1). The remaining individuals either

114   had haplotypes diagnostic of the Critically Endangered Española Island species ($n = 70$),

115   haplotypes shared between the Santiago Island and Volcano Wolf species ($n = 44$), or a

116   haplotype shared between the species from San Cristóbal and Santa Cruz Islands ($n = 1$). The

117   proportions of haplotypes associated with the Española and Floreana Island species reported here

118   (46% and 23%, respectively) are substantially higher than previously detected among individuals

119   from Volcano Wolf[25,26] (5% and 2%). This is likely due to our targeting of saddle-backed

120   individuals in 2015 versus sampling broadly during previous surveys[25–28].

121   Bayesian clustering analyses of microsatellite genotypic data using the method

122   implemented in STRUCTURE[36] revealed that 127 of the 150 tortoises sampled have ancestry

4

123 assignments (Q-values) to the extinct Floreana species (average Q-value ± SD = 0.87 ± 0.21;

124 range 0.16–0.99; Figure 2; Supplementary Figures S4-S6). Twenty-three individuals did not

125 show evidence of Floreana ancestry, being assigned to the two genetically distinct populations

126 (PBL and PBR) of *C. becki*, the endemic Volcano Wolf species. Of those individuals with

127 Floreana Q-values, 30 had Floreana mitochondrial haplotypes (Figure 2).

128 We conducted additional assignment tests using genotypes generated by simulated

129 matings within and among four possible parental lineages (PBL, PBR, Española Island, and

130 Floreana Island). Use of simulated individuals in assignment tests has been suggested to improve

131 accuracy and efficiency when distinguishing hybrid individuals[37]. Q-values from STRUCTURE

132 for the simulated hybrids were in the range expected for each "class" of hybrid (Supplementary

133 Figure S4), indicating we have the ability to identify individuals with Floreana ancestry across

134 various levels of admixture.

135 Analyses with the program GeneClass2 version 2.0[38] identified two tortoises with strong

136 assignment to the Floreana species, being classified as either purebreds or backcrosses between a

137 F_1 × purebred Floreana tortoise (Supplementary Table S2). This program also identified

138 additional 63 tortoises that were assigned to categories with Floreana ancestry. Further analysis

139 of these individuals found a large number of F_1 hybrids and backcrosses both when re-running

140 STRUCTURE (Q-values between 0.40–0.77; Supplementary Table S2), as well as when we used

141 NEWHYBRIDS version 1.1[39] (Table 1) and discriminant analysis of principal components

142 (DAPC ; Figure 3, Table 1). Admixture between the saddle-backed species from Floreana and

143 Española Islands (*n* = 43) was more common than between Floreana and either of the two

144 endemic domed *C. becki* populations (PBL and PBR; *n* = 22), indicative of positive assortative

145 mating between the two saddle-backed species.

146 A critical factor when planning a captive breeding program is having accurate ancestry

147 assignments. To examine the precision and accuracy of our ancestry assignments we conducted

148 additional analyses in STRUCTURE. First, jackknifing the loci to test robustness of Q-values to

149 reductions in the number of markers, and using the ANCESTDIST option in STRUCTURE,

150 which collects information about the distribution of Q-values for each individual in our dataset.

151 Jackknifing showed that individual Q-values were robust to reductions in the number of markers:

152 average variation in individual Q-values was 0.027 (SD=0.033, range <0.001-0.149;

153 Supplementary Figure S5). Second, use of the ANCESTDIST option highlighted that, as

154   expected, for hybrid individuals there is uncertainty around the specific estimate of Q-values

155   (range 0.002-0.484; Supplementary Figure S6). Inclusion of the simulated individual in our

156   STRUCTURE run led to a general decrease in the magnitude of Q-values observed

157   (Supplementary Table S2 and Supplementary Figure S7). Taken together, these analyses suggest

158   that Q-values should not be taken as direct measure of proportional ancestry, but that our

159   markers are powerful in detecting the presence of ancestry from the extinct Floreana species.

160   Neither mitochondrial DNA nor microsatellite data identified individuals with ancestry

161   from the Pinta Island species. However, a previous genetic estimate suggested that only 60–70

162   tortoises with Pinta ancestry are present on Volcano Wolf[25], whereas capture-mark-recapture

163   methods employed during our expedition estimated that a total of ~5,000–6,000 tortoises

164   occurred in the area searched. Moreover, despite our substantial search effort, we explored only

165   ~26% of the total tortoise-occupied range on the volcano (J. Gibbs unpublished data). Therefore,

166   it is possible that individuals with Pinta ancestry still live on Volcano Wolf, but went undetected

167   in 2015.

168   Of the 38 individuals currently housed in the captive breeding center, 23 tortoises (9

169   males and 14 females) were found to have nuclear ancestry from Floreana across multiple

170   assignment methods (Table 1) with Q-values from STRUCTURE ranging between 0.44–0.77,

171   when simulated individuals were also included in the analyses. In addition, 12 of these 23

172   individuals have mitochondrial haplotypes from the Floreana species (Supplementary Table S2).

173   Together, these 23 individuals now form the core of a genetically-informed captive breeding

174   program aimed at repatriation of tortoises to Floreana Island. For the purposes of the breeding

175   program, all 23 genetically important individuals will be included, regardless of mitochondrial

176   lineage, in order to capture and maintain as much nuclear genetic diversity from the Floreana

177   species as possible. The program is modeled after another one developed for the Española

178   species initiated with only 15 founders. Over 50 years, the Española tortoise program generated

179   >2000 repatriates with nearly 1,000 surviving tortoises now reproducing independently on their

180   native island[17,40,41].

181   A key attribute for success of such breeding programs is that the founding individuals are

182   unrelated, as high relatedness can lead to inbreeding depression. Accordingly, we examined

183   relatedness among the 23 individuals with Floreana ancestry in the breeding center. The analysis

184   showed that most individuals were unrelated (average Queller and Goodnight's[42] relatedness = -

185   0.04, range -0.58–0.63; Supplementary Figure 3). Although the 23 tortoises in captivity represent

186   a promising founding population, 44 tortoises identified during the 2015 expedition, but left on

187   Volcano Wolf, are also good candidates for the Floreana breeding and repatriation program

188   (Table 1). These individuals, once re-located on Volcano Wolf, could be incorporated into the

189   breeding program to further expand the genetic diversity of the founder population.

190   The Floreana tortoise breeding program will be designed to maximize founder

191   contributions and *C. elephantopus* genome representation in the resulting progeny, while

192   promoting *in situ* population growth and minimizing costs to the Galápagos National Park

193   Directorate. Depending on the goals and priorities of Park decision-makers, complete genome

194   recovery may not be reached before releases of offspring begin. However, the high proportion of

195   Floreana ancestry and low relatedness evident in the current breeding individuals indicate that

196   the 23 founding individuals and their resulting progeny will provide a good starting point for

197   restoring the species.

198   Our discovery raises the possibility that the extinct Floreana species could be revived. In

199   this case, tortoises with Floreana ancestry are living 'genomic archives' that retain the

200   evolutionary legacy of the extinct species, removing the need for the cloning methods that have

201   been proposed to bring back extinct species[43]. The Floreana tortoise breeding program is

202   anticipated to generate thousands of offspring over the next few decades. When repatriated to

203   Floreana Island, these tortoises can once again play their critical role as ecosystem engineers[19].

204   In addition, giant tortoises are a major tourist attraction in Galápagos[44]; tortoise restoration on

205   Floreana Island should create new economic opportunities for the island's few human residents.

206   Ironically, the opportunity to revive this "lost species" today was created by the same early

207   visitors to the archipelago whose activities imperiled most giant tortoise species and drove some

208   into outright extinction.

209

210   **Materials and Methods**

211   **Sampling and lab methods**

212   All samples were collected under CITES permit 15US209142/9, Galápagos Park Permit

213   PC-75-16, and in accordance with Yale Institutional Animal Care and Use Committee (IACUC)

214   permit number 2016-10825. Samples were collected over a 10-day expedition in November 2015

215   by a team of 36 researchers plus Galápagos National Park rangers. Groups of 3–4 searchers were

7

216   assigned distinct areas on Volcano Wolf, collectively totaling ~36 km^2. The search area was

217   chosen based on a previous survey of tortoises on Volcano Wolf, which indicated that

218   individuals with Pinta and Floreana ancestry were found to be in their highest densities on the

219   western slopes of Volcano Wolf[25]. For all tortoises encountered, sex, age, and GPS coordinates

220   were recorded; for saddle-backed individuals, photographs were taken, blood samples were

221   collected for DNA analysis, and each individual was injected with a passive integrated

222   transponder (PIT) tag under the skin for identification in the future. In cases where blood was

223   taken, ~2 ml of blood was collected from the brachial vein of one of the front legs of the tortoise

224   and preserved in a lysis buffer containing 0.1 M Tris buffer, 0.1 M EDTA, 0.2 m NaCl, and 1%

225   SDS, pH 8.0. All tortoises were uniquely marked with paint when first encountered. Starting on

226   the sixth day, we re-searched areas and recorded whether any individual had been previously

227   encountered for the purposes of a capture-mark-recapture estimate of population size. In total,

228   1,333 tortoises were encountered, of which 144 had saddle-backed morphology. The 32 tortoises

229   removed from Volcano Wolf were initially carried in nets by helicopter from the flanks of the

230   volcano to the expedition ship anchored in Puerto Bravo harbor, and then transported to the

231   Galápagos National Park captive tortoise breeding facility on Santa Cruz Island.

232   DNA was extracted from 150 blood samples using Qiagen blood and tissue extraction

233   kits. These samples included the 144 saddle-backed individuals mentioned above along with six

234   individuals already housed at the Galápagos National Park Breeding Center on Santa Cruz Island

235   that were previously identified to have Floreana ancestry[35]. All samples were sequenced at

236   ~700bp of the mitochondrial DNA control and were genotyped at 12 dinucleotide

237   microsatellites, using previously developed protocols (detailed procedures in Supplementary

238   Methods).

239   **Ancestry assignment**

240   For ancestry assignment based on mitochondrial DNA, the new sequences were aligned

241   to a reference database of 123 previously observed haplotypes, representing all extant and extinct

242   species[29]. Ancestry was assigned by determining shared haplotypes using the program TCS

243   version 1.21[45]. Evolutionary relationships were viewed with the program Network version 5.0

244   (fluxus-engineering.com; Supplementary Figure 1).

245   For ancestry assignment based on microsatellite genotypic data, we took a two-step

246   process using STRUCTURE version 2.3.4[36,46]. First, we ran STRUCTURE with a reference

247   dataset of 277 samples including all extant and extinct species to confirm the number of expected

248   genetic clusters (K) present in the archipelago (for full parameters see Supplementary Methods).

249   In this case, the optimal K was 12, which corresponds to previously described results[25,26]

250   (Supplementary Figure 2). Second, with archipelago-wide K established, we reduced the

251   reference dataset to the seven clusters previously found to be on Volcano Wolf[29]: Española, San

252   Cristóbal, Central Isabela (La Cazuela, Volcano Alcedo, and Volcano Darwin), Floreana, Pinta,

253   and the two genetically distinct endemic Volcano Wolf populations, PBL, and PBR. To this

254   reduced dataset (155 reference samples in total), the new samples were added and we set K to 7,

255   leaving all other parameters unchanged.

256   To further quantify the genetic ancestry of individuals, assignment tests were undertaken

257   using three additional methods. Prior to implementing these methods, we expanded the reference

258   database to include simulated genotypes. These genotypes were simulated using HYBRIDLAB

259   version 1.0[47], and corresponded to individuals arising from crosses within and among the

260   genetically distinct *C. becki* populations on Volcano Wolf (PBL and PBR), and the two saddle-

261   backed species from Española and Floreana (see Supplementary Methods). These four lineages

262   were chosen due to their high prevalence in the mitochondrial DNA-based assignments and their

263   contribution to microsatellite genotypes in admixed individuals. Simulated individuals

264   represented explicit ancestry categories: pure parental populations, first-generation (F_1) crosses,

265   second-generation (F_2) crosses, and backcrosses (i.e., matings between F_1's and their respective

266   purebred lineages).

267   With these new genotypes added into the reference set, we first used GeneClass version

268   2.0[38] to calculate the probability that an individual's genotype assigns to a population. Second,

269   we used NEWHYBRIDS version 1.1[39] to compute the posterior probability of various hybrid

270   classes for each individual. In this case, only four pairwise combinations of parental populations

271   were considered to focus on identifying potential hybrids involving Floreana ancestry: 1)

272   Española × Floreana, 2) PBL × Floreana, 3) Floreana × PBR, and 4) Española × PBR. Third, we

273   assigned ancestry using a multivariate approach, DAPC, as implemented in the R package

274   *adegenet* version 2.0.1[48,49]. In this case, the simulated genotypes from crossess among Floreana,

275   Española, PBL and PBR were initially used as a training dataset to define the principal

276   components and discriminant functions. The empirical genotypes for putative hybrids were then

277   transformed with principal components analysis (PCA) based on the centering and scaling of the

278   training set, and positioned onto the discriminant functions. The NEWHYBRIDS and DAPC

279   analyses were carried out only on individuals with ancestry from Floreana based on the results

280   from the GeneClass analyses (Supplementary Table S2). Finally, we re-ran STRUCTURE using

281   the simulated crosses as the reference populations.

282   Precision and accuracy of ancestry estimates were tested with two additional sets of

283   STRUCTURE analyses. First, we jackknifed our loci, sequentially removing one locus from the

284   dataset and rerunning STRUCTURE. Second, we used the ANCESTDIST option within

285   STRUCTURE. See Supplementary Materials for full methods.

286   **Relatedness Analysis**

287   For the 23 individuals in the breeding center with Floreana ancestry, we calculated

288   pairwise relatedness using the estimator of Queller and Goodnight[42]. This was the highest ranked

289   of eight tested relatedness estimators tested by the program irelr[50] which considers a composite

290   score that incorporates estimates of bias, variance, skewness, and kurtosis. The empirical

291   distribution of relatedness values was compared to the distributions of 10,000 simulated pairs of

292   individuals for each of four relatedness categories (unrelated, half sibs, full sibs, and parent–

293   offspring). The pairwise relatedness distribution of both the empirical and simulated data were

294   calculated using irelr[50].

295   **<u>References</u>**

296   1.   Dirzo, R. *et al.* Defaunation in the Anthropocene. *Science (80-.).* **345,** 401 LP-406 (2014).

297   2.   Novacek, M. J. & Cleland, E. E. The current biodiversity extinction event: Scenarios for

298   mitigation and recovery. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 5466–5470 (2001).

299   3.   Corlett, R. T. Restoration, Reintroduction, and Rewilding in a Changing World. *Trends*

300   *Ecol. Evol.* **31,** 453–462 (2016).

301   4.   Weeks, A. R. *et al.* Assessing the benefits and risks of translocations in changing

302   environments: a genetic perspective. *Evol. Appl.* **4,** 709–725 (2011).

303   5.   Canessa, S. *et al.* Planning for ex-situ conservation in the face of uncertainty. *Conserv.*

304   *Biol.* **30,** 599–609 (2016).

305   6.   Hoffmann, M. *et al.* The Impact of Conservation on the Status of the World's Vertebrates.

306   *Science (80-.).* **330,** 1503–1509 (2010).

307   7.   Doremus, H. & Pagel, J. E. Why Listing May Be Forever: Perspectives on Delisting under

308   the U.S. Endangered Species Act. *Conserv. Biol.* **15,** 1258–1268 (2001).

309 8.   Taylor, M. F. J., Suckling, K. F. & Rachlinski, J. J. The Effectiveness of the Endangered
310   Species Act: A Quantitative Analysis. *Bioscience* **55,** 360–367 (2005).
311 9.   Sherkow, J. S. & Greely, H. T. What If Extinction Is Not Forever? *Science (80-.).* **340,**
312   32–33 (2013).
313 10. Cohen, S. The Ethics of De-Extinction. *Nanoethics* **8,** 165–178 (2014).
314 11. Jones, K. E. From dinosaurs to dodos: who could and should we de-extinct? *Front.*
315   *Biogeogr.* **6,** (2014).
316 12. Phillips, R. B., Wiedenfeld, D. A. & Snell, H. L. Current status of alien vertebrates in the
317   Galápagos Islands: invasion history, distribution, and potential impacts. *Biol. Invasions*
318   **14,** 461–480 (2012).
319 13. Guézou, A. *et al.* An Extensive Alien Plant Inventory from the Inhabited Areas of
320   Galapagos. *PLoS One* **5,** e10276 (2010).
321 14. Harper, G. A. & Carrion, V. Introduced rodents in the Galapagos: colonisation, removal
322   and the future. *Isl. Invasives Erad. Manag.* 63–66 (2011).
323 15. Campbell, K., Donlan, C. J., Cruz, F. & Carrion, V. Eradication of feral goats Capra
324   hircus from Pinta Island, Galápagos, Ecuador. *Oryx* **38,** 328–333 (2004).
325 16. Fabiani, A. *et al.* Conservation of Galápagos land iguanas: genetic monitoring and
326   predictions of a long-term program on the island of Santa Cruz. *Anim. Conserv.* **14,** 419–
327   429 (2011).
328 17. Milinkovitch, M. C. *et al.* Genetic analysis of a successful repatriation programme: giant
329   Galapagos tortoises. *Proc. R. Soc. B-Biological Sci.* **271,** 341–345 (2004).
330 18. Benitez-Capistros, F., Hugé, J., Dahdouh-Guebas, F. & Koedam, N. Exploring
331   conservation discourses in the Galapagos Islands: A case study of the Galapagos giant
332   tortoises. *Ambio* **45,** 706–724 (2016).
333 19. Gibbs, J. P., Sterling, E. J. & Zabala, F. J. Giant Tortoises as Ecological Engineers: A
334   Long-term Quasi-experiment in the Galápagos Islands. *Biotropica* **42,** 208–214 (2010).
335 20. Poulakakis, N. *et al.* Description of a New Galapagos Giant Tortoise Species
336   (Chelonoidis; Testudines: Testudinidae) from Cerro Fatal on Santa Cruz Island. *PLoS One*
337   **10,** e0138779 (2015).
338 21. Márquez, C. *et al.* Estado actual de las poblaciones de tortugas terrestres gigantes
339   (Geochelone spp., Chelonia: Testudinae) en las islas Galápagos. *Ecol. Apl.* **3,** (2004).

340 22. Porter, D. *Journal of a cruise made to the Pacific Ocean*. (Wiley & Halsted, 1822).

341 23. Townsend, C. H. The Galápagos tortoises in relation to the whaling industry: A study of
342   old logbooks. *Zoologica* **4,** 55–135 (1925).

343 24. IUCN. The IUCN Red List of Threatened Species. Version 2016-2. (2016). Available at:
344   www.iucnredlist.org. (Accessed: 10th November 2016)

345 25. Edwards, D. L. *et al.* The genetic legacy of Lonesome George survives: Giant tortoises
346   with Pinta Island ancestry identified in Galapagos. *Biol. Conserv.* **157,** 225–228 (2013).

347 26. Garrick, R. C. *et al.* Genetic rediscovery of an 'extinct' Galapagos giant tortoise species.
348   *Curr. Biol.* **22,** R10–R11 (2012).

349 27. Poulakakis, N. *et al.* Historical DNA analysis reveals living descendants of an extinct
350   species of Galapagos tortoise. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 15464–15469 (2008).

351 28. Russello, M. A. *et al.* Lonesome George is not alone among Galápagos tortoises. *Curr.*
352   *Biol.* **17,** R317–R318 (2007).

353 29. Poulakakis, N., Russello, M., Geist, D. & Caccone, A. Unravelling the peculiarities of
354   island life: vicariance, dispersal and the diversification of the extinct and extant giant
355   Galapagos tortoises. *Mol. Ecol.* **21,** 160–173 (2012).

356 30. Benavides, E. *et al.* Lineage Identification and Genealogical Relationships Among
357   Captive Galapagos Tortoises. *Zoo Biol.* **31,** 107–120 (2012).

358 31. Garrick, R. C. *et al.* Lineage fusion in Galapagos giant tortoises. *Mol. Ecol.* **23,** 5276–
359   5290 (2014).

360 32. The Galapagos Conservancy. The Wolf Expedition: Resurrecting the extinct tortoise
361   species from Floreana and Pinta. (2015). Available at:
362   http://www.galapagos.org/newsroom/wolf-tortoise-search-outcome/.

363 33. Nicholls, H. Lonesome George 'look-alike' captured for tortoise-breeding effort. *Nature*
364   (2015). doi:10.1038/nature.2015.18955

365 34. Blakeslee, S. Scientists Hope to Bring a Galápagos Tortoise Species Back to Life. *The*
366   *New York Times* (2015).

367 35. Russello, M. A. *et al.* DNA from the Past Informs Ex Situ Conservation for the Future: An
368   "Extinct'' Species of Galapagos Tortoise Identified in Captivity. *PLoS One* **5,** (2010).

369 36. Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using
370   multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164,**

371   1567–1587 (2003).

372 37. Vähä, J.-P. & Primmer, C. R. Efficiency of model-based Bayesian methods for detecting

373   hybrid individuals under different hybridization scenarios and with different numbers of

374   loci. *Mol. Ecol.* **15,** 63–72 (2006).

375 38. Piry, S. *et al.* GENECLASS2: A Software for Genetic Assignment and First-Generation

376   Migrant Detection. *J. Hered.* **95,** 536–539 (2004).

377 39. Anderson, E. C. & Thompson, E. A. A model-based method for identifying species

378   hybrids using multilocus genetic data. *Genetics* **160,** 1217–1229 (2002).

379 40. Milinkovitch, M. *et al.* Giant Galapagos tortoises; molecular genetic analyses identify a

380   trans-island hybrid in a repatriation program of an endangered taxon. *BMC Ecol.* **7,**

381   doi:10.1186/1472-6785-7-2 (2007).

382 41. Gibbs, J. P., Hunter, E. A., Shoemaker, K. T., Tapia, W. H. & Cayot, L. J. Demographic

383   Outcomes and Ecosystem Implications of Giant Tortoise Reintroduction to Española

384   Island, Galapagos. *PLoS One* **9,** e110742 (2014).

385 42. Queller, D. C. & Goodnight, K. F. Estimating relatedness using genetic-markers.

386   *Evolution (N. Y).* **43,** 258–275 (1989).

387 43. Shapiro, B. Mammoth 2.0: will genome engineering resurrect extinct species? *Genome*

388   *Biol.* **16,** 228 (2015).

389 44. Medina, C., Aravena, C. & Vásquez, F. *Economic valuation of the giant tortoise*

390   *conservation in Galapagos Islands Carlos*. (2012).

391 45. Clement, M., Posada, D. & Crandall, K. A. TCS: a computer program to estimate gene

392   genealogies. *Mol. Ecol.* **9,** 1657–1659 (2000).

393 46. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using

394   multilocus genotype data. *Genetics* **155,** 945–959 (2000).

395 47. Nielsen, E. E., Bach, L. A. & Kotlicki, P. HYBRIDLAB (version 1.0): a program for

396   generating simulated hybrids from population samples. *Mol. Ecol. Notes* **6,** 971–973

397   (2006).

398 48. Jombart, T. adegenet: a R package for the multivariate analysis of genetic markers.

399   *Bioinformatics* **24,** 1403–1405 (2008).

400 49. Jombart, T. & Ahmed, I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP

401   data. *Bioinformatics* **27,** 3070–3071 (2011).

402 50. Gonçalves da Silva, A. & Russello, M. A. iRel: software for implementing pairwise
403   relatedness estimators and evaluating their performance. *Conserv. Genet. Resour.* **3,** 69–
404   71 (2011).
405 51. R Core Team. R: A language and environment for statistical computing, reference index
406   version 3.2.2. (2015).
407 52. Hijmas, R. J. raster: Geographic Data Analysis and Modeling version 2.5-8. (2016).
408

409 **Acknowledgements**

422 **Author Contributions**

423 JMM, NP, LB, CC, EAH, WT, DR, JC, JPG and AC assisted with field logistics and sample
424 collection. JMM, MCQ, and AV conducted laboratory analyses. JMM, MCQ, and NP conducted
425 data analyses. LBB, RCG, MAR, and DLE provided technical guidance and conceptual advice
426 for data analyses. JMM drafted the manuscript with assistance from MCQ, LBB, MAR, JPG, and
427 AC. JMM, MCQ, NP, LBB, RCG, MAR, CC, DLE, EAH, WT, JPG, and AC discussed the
428 results and implications and commented on the manuscript at all stages.
429

430 **Additional Information**

431 Microsatellite genotypes and mitochondrial sequences are available on a server hosted by the
432 College of Environmental Science & Forestry, State University of New York, Syracuse

14

433   (http://www.esf.edu/efb/gibbs/Miller_et_al_data_archive.zip).   The   authors   have   no   financial

434   conflicts of interest. Correspondence and requests for materials should be addressed to either

435   joshua.miller@yale.edu or adalgisa.caccone@yale.edu.

436

437 **Figures**
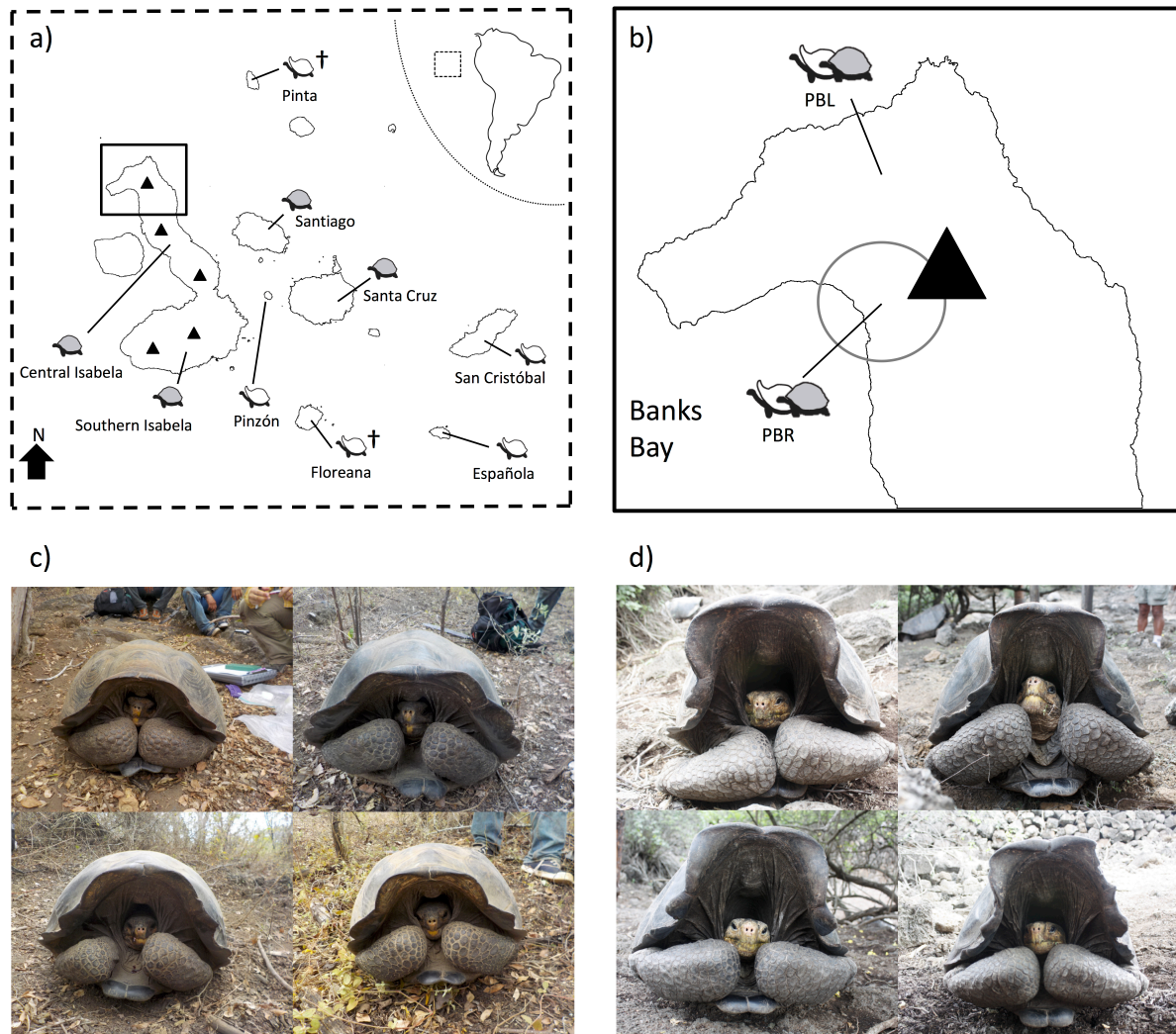
438

439 Figure 1: Distribution of tortoises among Galápagos Islands and representative photos of tortoise
440   carapace morphology. A) Map of the distribution of tortoises among Galápagos Islands
441   along with cartoons indicating carapace morphology for each. Light grey shading
442   indicates domed morphology, unshaded indicates saddle-backed. Extinct species are noted
443   with †. B) Larger view of Volcano Wolf on northern Isabela Island. The circle indicates
444   the approximate field location of the current study. Examples of Galápagos giant tortoises
445   with domed (C) saddle-backed (D) morphology. Maps were created with R (version
446   3.2.4[51]) using the raster package (version 2.5-8[52]).

447
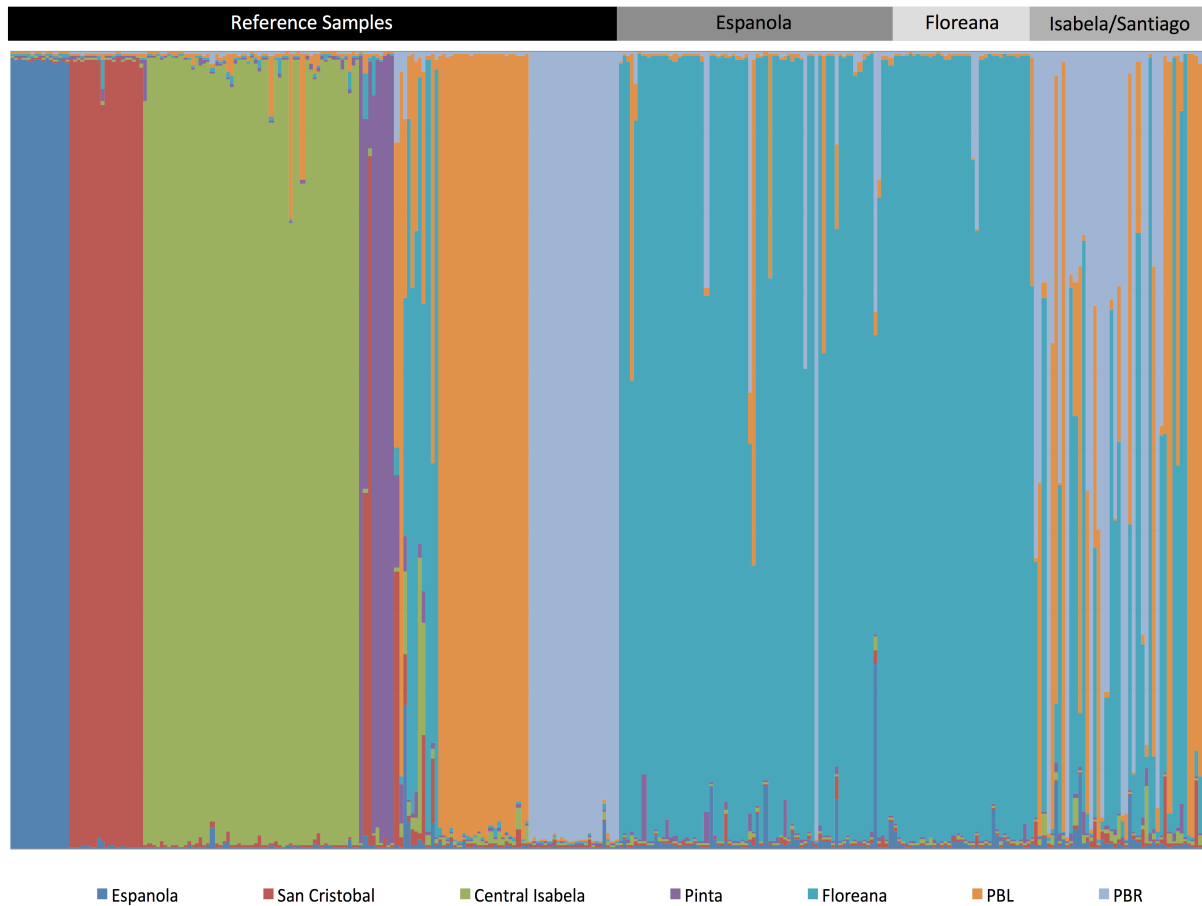
Figure 2: STRUCTURE[36,46] plot for 155 reference samples along with the 150 newly collected individuals. Each individual is represented as a vertical bar, with colors denoting the different genetic clusters, as indicted below. The proportion of color in a bar is equal to the ancestry (Q-value) to a given cluster. Reference samples are those under the black horizontal bar, the newly collected samples are those under the grey horizontal bar with shades and labels corresponding to mitochondrial lineage (the final sample on the far right has a haplotype shared by the Islands of San Cristóbal and Santa Cruz).
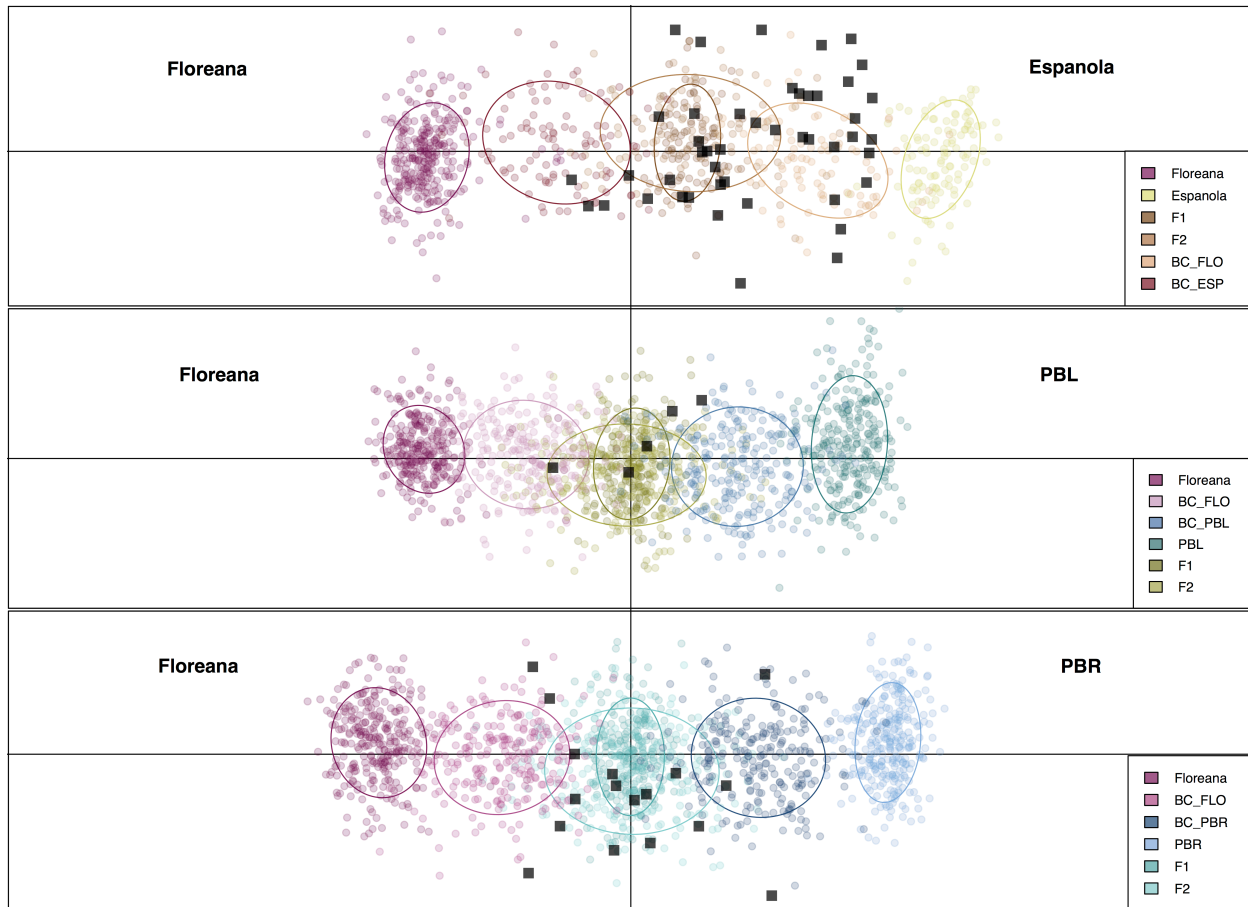
Figure 3: Scatterplot of the first two principal components of DAPC showing assignment of unknown individuals to ancestry categories. The rows show pairwise combinations of parental populations. Top: Española × Floreana; middle: Floreana × PBL; bottom: Floreana × PBR. Parental populations are positioned at the ends of each plot with color coded first generation (F_1), second generation (F_2), and each backcross (BC) individuals between them. Ellipses encompass ~67% of the cloud of points for each group. These genotypes were used as a training set to define the discriminant functions and optimal number of PCs to retain. The Volcano Wolf tortoises with mixed ancestry are identified by black squares where placement is based on the previously defined discriminant functions.

466 Table 1: Number of individuals assigned to ancestry categories by DAPC (A) and
467 NEWHYBRIDS (B)

468 A)

Location		BC to ESP	BC to FLO	BC to PBL	BC to PBR	FLO/ESP F_1 or F_2	FLO/PBL F_1 or F_2	FLO/PBR F_1 or F_2	Total
	PNG	9	2	0	0	8	1	3	23
	VW	13	5	1	3	12	2	8	44
Total		22	7	1	3	20	3	11	67

469

470 B)

Location		BC to ESP	BC to FLO	BC to PBL	BC to PBR	FLO/ESP F_1 or F_2	FLO/PBL F_1 or F_2	FLO/PBR F_1 or F_2	Total
	PNG	0	12	0	0	4	1	2	19
	VW	0	7	0	0	17	4	8	36
Total		0	19	0	0	21	5	10	55

471

472 The current location of individuals is listed, as Galápagos National Park Breeding Center (PNG)
473 or Volcano Wolf (VW). Abbreviations are: BC = backcross, F_1 = first generation mating, F_2 =
474 second generation mating (F_1 x F_1), ESP = Española, and FLO = Floreana. Note that the total
475 number of samples assigned differs between the two methods as some individuals did not exceed
476 the posterior probability threshold for classification in NEWHYBRIDS (see Supplementary
477 Methods).