PNAS
www.pnas.org

Supplementary Information for

Triad hybridization via a conduit species

Peter R. Grant and B. Rosemary Grant

Corresponding author

Peter R. Grant

Email: prgrant@princeton.edu

This PDF file includes:

- Supplementary text
- Figs. S1 to S2
- Table S1
- Captions for datasets S1 to S5
- References for SI reference citations

Other supplementary materials for this manuscript include the following:

- Datasets S1 to S5

Triad hybridization via a conduit species

Peter R. Grant and B. Rosemary Grant

Section 1 The problem of identifying species, hybrids and backcrosses

We began a field study of finches on Daphne Major in 1973. Details of how we recognized and classified the individuals to species are given in reference 1. *G. fortis* and *G. scandens* were easy to distinguish by their blunt or pointed beaks. We captured a few individuals of *G. fuliginosa* as well, probably immigrants from Santa Cruz island situated approximately 8 km to the south. They were morphologically similar to the smallest *G. fortis* on Daphne. To classify them we used the upper limits to beak size distributions in a large sample on nearby Santa Cruz island. These limits were considered as a boundary demarcating them from *G. fortis*, which are distinctly larger than *G. fuliginosa* on Santa Cruz island. Then, in 1976, we began a breeding study of the finches on Daphne (2). We assumed that all individuals were members of one of the three species, and none were hybrids, as explained on p. 130 in reference 1:

"The way in which inbreeding is studied shows how this problem should be addressed. Estimated inbreeding coefficients have meaning only in relation to a base population. The base population, that is the population at the start of a

study, is likely to contain an unknown number of inbred individuals and an unknown number of breeding pairs of related individuals. Nevertheless, it has to be assumed that members of all mated pairs in the starting (Fo) generation are unrelated (Falconer 1989). The same principle and practice hold for field studies of hybridization when hybrid individuals cannot be unambiguously recognized from their phenotypes (or genotypes) as here. All individuals in the starting generation are assumed to belong to one species or another, and hybrids are assumed to be absent."

We found a few instances of interbreeding between *G. fortis* and the other two species, raising the possibility that some of the breeders were not pure members of the three species but actually hybrids. However, although apparent hybrids were produced, they all died either that year or the next one during a severe drought. Only later, after the El Niño event of 1983 had transformed the vegetation, did we discover that hybrids could survive long enough to breed (3). This finding put a different complexion on the initial study of breeding as it raised again the question of whether some of the initial breeders were in fact hybrids. In 1988 we began collecting small samples of blood for DNA analysis of every bird we handled, and with assignment tests applied to patterns of microsatellite variation (see next section) we were able to examine the question of hybrids (4). The samples included some old birds. The oldest were two *G. scandens* that had hatched in 1975. Further, we were able to check for extra-pair paternity. We discovered that some finches of each of the three species were admixtures of the genomes of two or more species. Some of the assignments confirmed our identifications from pedigrees, whereas others did not. Therefore, some of the

initial breeders that were not genotyped could plausibly have been hybrids, meaning F_1s and backcrosses. As a consequence, individuals identified as species may in fact have been hybrids, and our classification of particular hybrids from pedigree data may have been erroneous if their parents were not pure species but hybrids. Unfortunately, we had no genetic data from the initial study and therefore, in the absence of information to the contrary, we had to assume that all breeders were pure members of the respective (phenotypic) species. Subsequent analyses confirmed the presence of hybrids in our earlier samples (Table 1 main text), and identified extra-pair mating as one way in which hybrids were produced (5)

Section 2 The genetic signal of microsatellite-based assignments

In a previous paper on the finches of Daphne Major island (4) we used alleles at fourteen unlinked, polymorphic, autosomal loci and the program STRUCTURE to assign individuals to species and hybrids (admixtures). Species were initially identified by morphology, and then individuals were recognized as members of one of the three species when the probability of assignment to a species reached or exceeded a threshold of 0.90. If the 0.90 criterion was not met they were assigned to hybrids. The value is arbitrary, although often used (6-9), and was chosen as a compromise between precision and inclusiveness. The choice recognizes that markers are not diagnostic, and genotyping errors add further uncertainty. Separating individuals into classes (F_1 and backcrosses) was found to be unreliable. McFarlane and Pemberton in reference 10 refer to choosing a threshold for delimiting species and hybrids as a balancing act between type I

errors and type II errors --- mistakenly assigning parental species individuals to a hybrid and assigning a hybrid individual to a parental species. Advanced backcrosses will be treated as parental species if a high level of type II errors is accepted, as may be the case with 0.90. Thus, the 0.90 threshold is conservative in the sense of probably underestimating the true frequency of admixtures.

In the main text we use the 14 autosomal loci plus two sex-linked loci because sex ratios are not likely to be biased with large samples. Long-standing coexistence with hybridization suggests long-standing backcrossing. Evidence for this is the high frequency of assignment probabilities in the range 0.90 to 0.99 (Fig. S1), in both species, and the increase in these frequencies through time, in agreement with other evidence presented in the main text (Fig. 2). Following reference 10 we used the range of 0.37 to 0.63 probabilities as an approximate guide to the expected frequencies of F_1 hybrids. They constitute only 2.3 percent (N = 74) of the total 3165 genotyped individuals, or 19.0 percent of the 390 hybrids. Corresponding figures for the backcrosses (P = 0.64-0.89) are 8.2 percent of the total or 66.2 percent of the hybrids. The remainder, trihybrids, are 1.8 percent of the total or 14.9 percent of the hybrids.

Section 3 Concordance of phenotypic and genetic classifications

The degree to which classifications based on morphology and microsatellites are congruent provides a measure of consistency, reliability and usefulness of the classification. Ninety-three percent of phenotypic *G. fortis* (N = 1195) were classified as *G. fortis* by microsatellite assignments. A similar proportion (92 percent) of measured *G. scandens* individuals (N = 391) were classified as *G.*

scandens by both methods. Most of the remaining *G. fortis* (44/85) were classified as Ff and most of the remaining *G. scandens* (19/33) were classified as SF hybrids by microsatellites. The classifications were less consistent with *G. fuliginosa*. Sixty-one percent of 26 individuals were classified by both methods as *G. fuliginosa*, and most of the remainder (6/9) were assigned to Ff hybrids by microsatellites. Classification is least consistent with hybrids. Among pedigree-assigned Ff hybrids identified in the pedigrees, 0.46 were assigned to admixtures by microsatellites and the remainder to *G. fortis*. Among the FS hybrids identified in the pedigrees, 0.27 were assigned by microsatellites to admixtures and the remainder to *G. fortis* or *G. scandens*. Some hybrid identities were wrongly determined from the pedigrees because interbreeding individuals were incorrectly assumed to be non-hybrids in the absence of information to the contrary (*SI Appendix,* section 1).

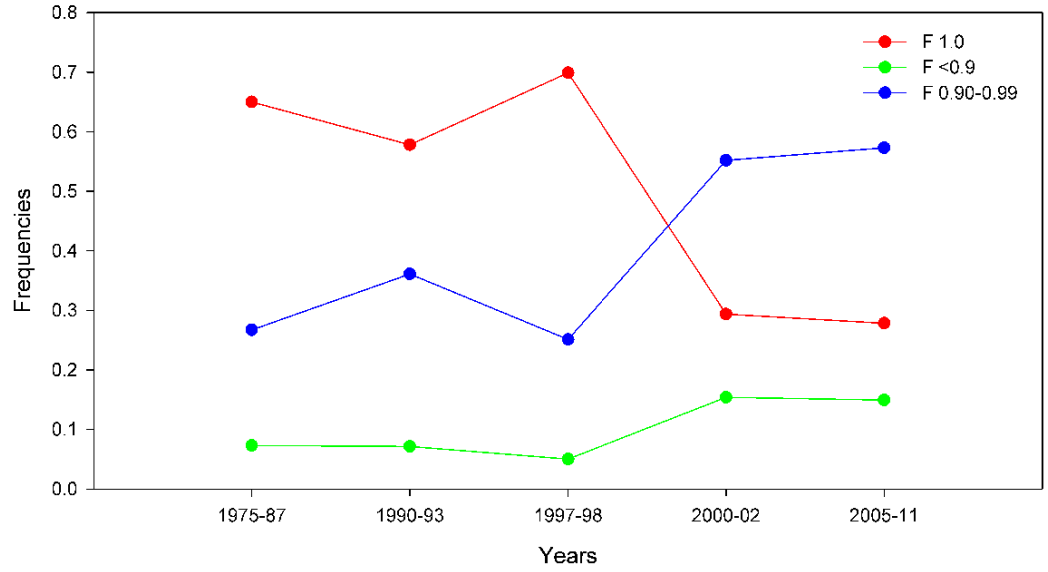
Another measure of reliability is the degree to which parents and offspring are classified in the same way by phenotype and microsatellite assignments. Consistency reflects inheritance of the defining allelic compositions of the parents. Ninety-seven percent of genotypically *G. fortis* x *G. fortis* parents (*N* = 598) produced only *G. fortis* offspring, and ninety-nine percent of genotypically *G. scandens* x *G. scandens* parents (*N* = 174) produced only *G. scandens* offspring. The few remaining pairs produced admixed offspring as a result of random segregation of parental genes in combinations that have affinities to another species.

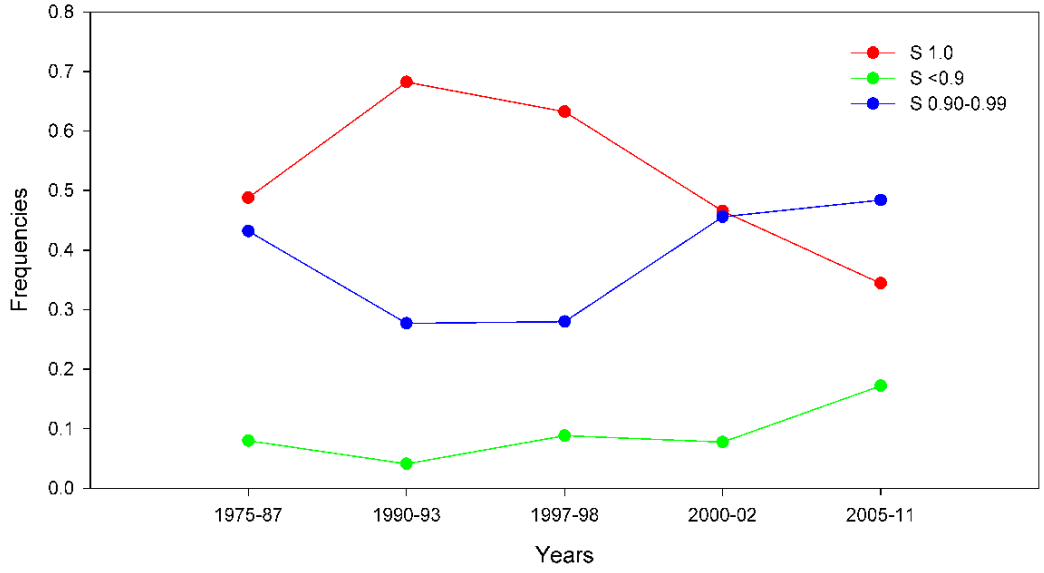
Section 4 Morphology

We used principal components analysis (PCA) to compare hybrids and species. We used morphological criteria to classify individuals to the three species in view of difficulties in distinguishing *G. fuliginosa* from Ff hybrids. We restricted the analysis of morphological variation among species to the period 1975-98, before introgression had proceeded far. However, we included all individuals identified as hybrids by microstallites across the whole period, 1978-2012. The total number of individuals was 1225. PC1 from an analysis of weight, wing length and tarsus length provides a measure of body size (Table S1). Loadings of all three measures are large, almost equal and positive. PC1 from an analysis of beak length, depth and width provides a measure of beak size, and again all loadings are large and positive. PC2 is a measure of beak pointedness or beak shape --- pointedness or bluntness: the loading of beak length is opposite in sign to the loadings of depth and width.

Section 5 Fitness

Di- and trihybrids differ in survival from the parental species to a small extent (main text). When the two sets of hybrids are combined with the species that gave rise to them, according to maximum probabilities of assignment, the difference between hybrids and species disappears in both the *G. fortis* group (main text, Fig. 5) and the *G. scandens* group (Fig. S2).


0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Frequencies
F 1.0
F <0.9
F 0.90-0.99
1975-87
1990-93
1997-98
2000-02
2005-11
Years


0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Frequencies
S 1.0
S <0.9
S 0.90-0.99
1975-87
1990-93
1997-98
2000-02
2005-11
Years

1

2

3

Fig. S1. Frequencies of individuals assigned to *G. fortis* (above) and *G. scandens* (below), with probabilities of 1.0 (red), 0.90-0.99 (blue) and 0.50 -0.90 (green). Sample sizes are 2182 *G. fortis* (F) and 1114 *G. scandens* (S).

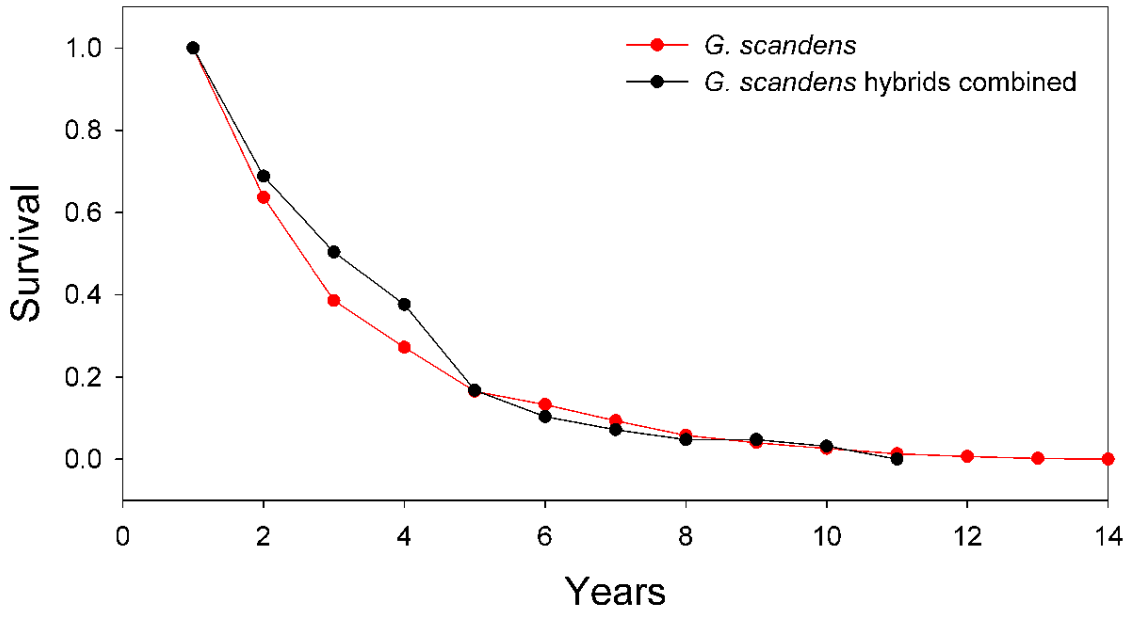


Fig. S2. Survival of the combined groups of *G. scandens* hybrids is almost identical to the survival of *G. scandens*.

Table S1. Results of two separate principal components analyses of body size and beak traits of 1225 finches. Numbers are loadings of the original variables on the principal components axes.

	PC1 body	PC1 beak size	PC2 beak shape
Percent variance	83.3	73.5	22.8
Weight	0.9290	.	.
Wing	0.9081	.	.
Tarsus	0.9013	.	.
Beak length	.	0.6764	0.7364
Beak depth	.	0.9300	-0.2861
Beak width	.	0.9396	-0.2471

Captions for datasets S1 to S5

Dataset S1. Dataset for Fig. 2. Frequency of hybrids produced each year of breeding. Hybrids are expressed as frequencies of the offspring produced in each of the *G. fortis* groups (*G. fortis*, Ff, FS) and *G. scandens* groups (*G. scandens*, SF, Sf, fS).

Dataset S2. Dataset for Fig. 3. Genetic convergence across five periods (1975-87, 1990-93, 1997-98, 2000-02, 2005-2011) separated by years of little or no breeding.

Dataset S3. Dataset for Fig. 4 and Tables 3 and 4. Morphological measurements of *G. fuliginosa*, *G. fortis* and *G. scandens* and hybrids.

Dataset S4. Dataset for Fig. 5 and S2. Survival of di- and trihybrids in relation to *G. fortis* and *G. scandens* from year 1 in the years 1987-2010.

Dataset S5. Dataset for Fig. S1. Frequencies of individuals assigned to *G. fortis* and *G. scandens* with probabilities of 1.0, 0.90-0.99 and 0.50 -0.90. Sample sizes are 2182 *G. fortis* and 1114 *G. scandens*.

References

1. P. R. Grant, Hybridization of Darwin's finches on Isla Daphne Major, Galápagos. *Philos. Trans. R. Soc. B* **340,** 127-139 (1993).
2. P. T. Boag, P. R. Grant, Darwin's Finches (*Geospiz*a) on Daphne Major, Galápagos: breeding and feeding ecology in a climatically variable environment. *Ecol. Monogr*. **54**, 463-489 (1984).

3. H. L. Gibbs, P. R. Grant, J. Weiland, Breeding of Darwin's Finches at an unusually early age in an El Niño year. *Auk* **101**, 872-874 (1984).
4. P. R. Grant, B. R. Grant, Conspecific versus heterospecific gene exchange between populations of Darwin's finches. *Philos. Trans. R. Soc. B* **365** 1065-1076 (2010).
5. P. R. Grant, B. R., Adult sex ratio influences mate choice in Darwin's finches. *Proc. Natl. Acad. Sci. USA* **116**, 12373-12382 (2019)
6. M. R. Kronforst, L. G. Young, L. M. Blume, L. E. Gilbert, Multilocus analyses of admixtures and introgression among hybridizing butterflies. *Evolution* **60**, 1254-1268 (2006).
7. M. L. Lancaster, N. J. Gemmell, S. Negro, S. Goldworthy, Ménage à trois on Macquarie Island: hybridization among three species of fur seal (*Arctocephalus* spp.) following historical population extinction. *Mol. Ecol.* **15**, 3681-3692 (2006).
8. J. Fogelqvist, A. V. Verkhozina, A. I. Katyshev, P. Pucholt, C. Dixelius, A. C. Rönneberg-Wästljung, M. Lascoux, S. Berlin, Genetic and morphological evidence of introgression between three species of willows. *BMC Evol. Biol*. **15**, 193 (2015).
9. M. L. Haines, J. Melville, J. Sumner, N. Clemann, D. G. Chapple, D. Stuart-Fox, Geographic variation in hybridization and ecological differentiation between three syntopic, morphologically similar species of montane lizards. *Mol. Ecol*. **25**, 2887-2903 (2016).
10. S. E. McFarlane, J. M. Pemberton, Detecting the true extent of introgression during anthropogenic hybridization. *Trends Ecol. Evol*. **34**, 315-326 (2019).

1