

Evolution of Darwin's finches and their beaks revealed by genome sequencing

Sangeet Lamichhaney^{1*}, Jonas Berglund^{1*}, Markus Sällman Almén¹, Khurram Maqbool², Manfred Grabherr¹, Alvaro Martinez-Barrio¹, Marta Promerová¹, Carl-Johan Rubin¹, Chao Wang¹, Neda Zamani^{1,3}, B. Rosemary Grant⁴, Peter R. Grant⁴, Matthew T. Webster¹ & Leif Andersson^{1,2,5}

Darwin's finches, inhabiting the Galápagos archipelago and Cocos Island, constitute an iconic model for studies of speciation and adaptive evolution. Here we report the results of whole-genome re-sequencing of 120 individuals representing all of the Darwin's finch species and two close relatives. Phylogenetic analysis reveals important discrepancies with the phenotype-based taxonomy. We find extensive evidence for interspecific gene flow throughout the radiation. Hybridization has given rise to species of mixed ancestry. A 240 kilobase haplotype encompassing the *ALX1* gene that encodes a transcription factor affecting craniofacial development is strongly associated with beak shape diversity across Darwin's finch species as well as within the medium ground finch (*Geospiza fortis*), a species that has undergone rapid evolution of beak shape in response to environmental changes. The *ALX1* haplotype has contributed to diversification of beak shapes among the Darwin's finches and, thereby, to an expanded utilization of food resources.

Adaptive radiations are particularly informative for understanding the ecological and genetic basis of biodiversity^{1,2}. Those causes are best identified in young radiations, as they represent the early stages of diversification when phenotypic transitions between species are small and interpretable and extinctions are likely to be minimal³. Darwin's finches are a classic example of such a young adaptive radiation^{3,4}. They have diversified in beak sizes and shapes, feeding habits and diets in adapting to different food resources^{4,5} (Extended Data Table 1). The radiation is entirely intact, unlike most other radiations, none of the species having become extinct as a result of human activities⁴.

Fourteen of the currently recognized species evolved from a common ancestor in the Galápagos archipelago (Fig. 1a) in the past 1.5 million years according to mitochondrial DNA (mtDNA) dating⁶; a fifteenth species inhabits Cocos Island. The radiation proceeded rapidly as a result of strong isolation from the South American continent, generation of new islands by volcanic activity, climatic oscillations caused by the El Niño phenomenon, and sea level changes associated with glacial and interglacial cycles over the past million years that led to repeated alternations of island formation and coalescence^{7,8}.

Traditional taxonomy of Darwin's finches is based on morphology³, and has been largely supported by observations of breeding birds^{4,5} and genetic analysis^{6,9}. However, the branching order of several recently diverged taxa is unresolved⁶ and genetic analysis of phylogeny has been limited to mtDNA and a few microsatellite loci. Some candidate genes for beak development are differentially expressed in species with different beak morphologies^{10–12}, but the loci controlling genetic variation in beak diversity among Darwin's finches remain to be discovered.

Here we report results from whole genome re-sequencing of 120 individuals representing all Darwin's finch species and two closely related tanagers, *Tiaris bicolor* and *Loxigilla noctis*¹³. For some species we collected samples from multiple islands (Fig. 1a). We comprehensively analyse patterns of intra- and interspecific genome diversity and phylogenetic relationships among species. We find widespread evidence of interspecific gene flow that may have enhanced evolutionary

diversification throughout phylogeny, and report the discovery of a locus with a major effect on beak shape.

Considerable nucleotide diversity

We generated approximately 10× sequence coverage per individual bird using 2 × 100 base-pair (bp) paired-end reads (Extended Data Fig. 1). Reads were aligned to the genome assembly of a female medium ground finch (*G. fortis*)¹⁴. We identified Z- and W-linked scaffolds on the basis of significant differences in read depth between males (ZZ) and females (ZW) (Supplementary Table 1) and generated a *G. fortis* mtDNA sequence through a combined bioinformatics and experimental approach. Stringent variant calling revealed approximately 45 million variable sites within or between populations. We found a considerable amount of genetic diversity within each population, in the range 0.3×10^{-3} to 2.2×10^{-3} (Extended Data Table 2), similar to that reported in other bird populations¹⁵ including island populations of the zebra finch¹⁶. We used these estimates of diversity to estimate effective population sizes of Darwin's finch species within a range of 6,000–60,000 (Supplementary Text). Extensive sharing of genetic variation among populations was evident, particularly among ground and tree finches, with almost no fixed differences between species in each group (Extended Data Fig. 2).

Genome-based phylogeny

According to the classical taxonomy of Darwin's finches, supported by morphological and mitochondrial (cytochrome b) data, warbler finches were the first to branch off, and ground and tree finches constitute the most recent major split^{3,6,9}. Our maximum-likelihood phylogenetic tree based on autosomal genome sequences is generally consistent with current taxonomy, but shows several interesting deviations (Fig. 1b). First, *Geospiza difficilis* occurring on six different islands forms a polyphyletic group separated into three distinct groups: (1) populations occupying the highlands of Pinta, Santiago and Fernandina, (2) populations occupying the low islands of Wolf and Darwin in the northwest^{3,6,9} and (3) the population on Genovesa in the northeast. This is consistent with

¹Department of Medical Biochemistry and Microbiology, Uppsala University, SE-751 23 Uppsala, Sweden. ²Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, SE-75007 Uppsala, Sweden. ³Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden. ⁴Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544, USA. ⁵Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, Texas 77843-4458, USA.

*These authors contributed equally to this work.

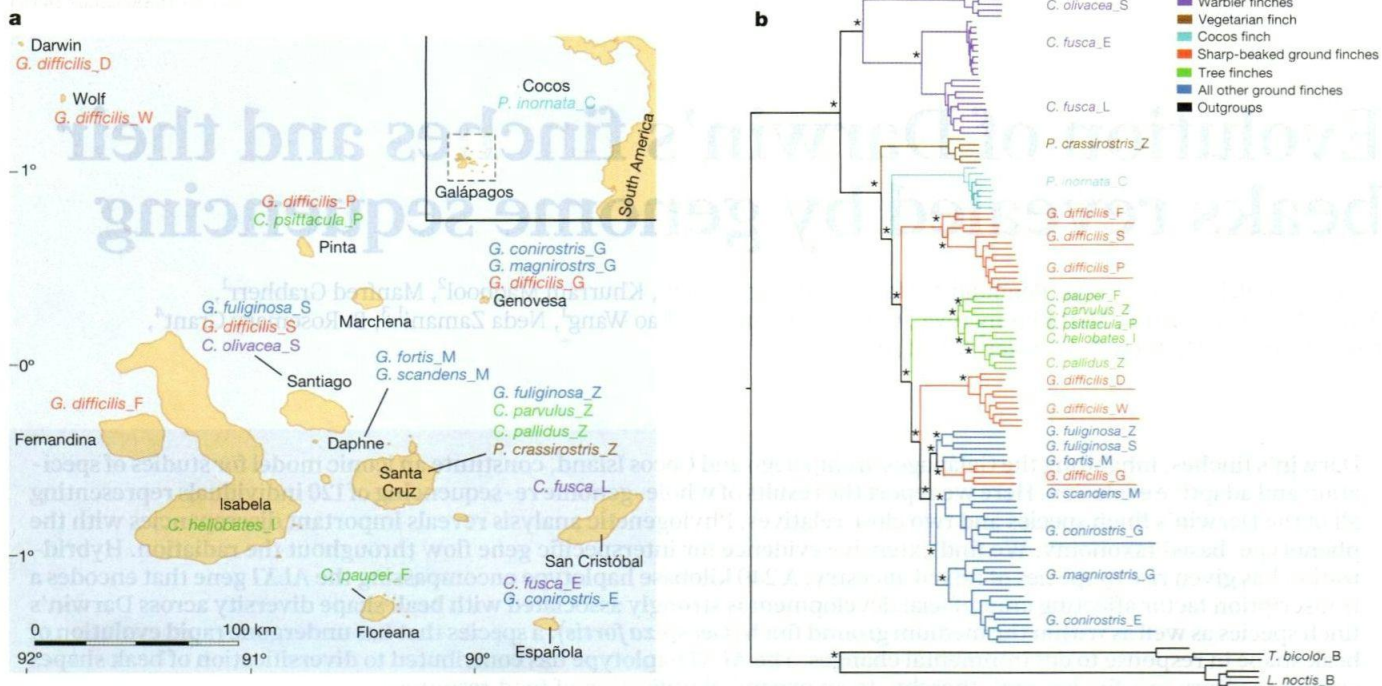


Figure 1 | Sample locations and phylogeny of Darwin's finches.

a, Geographical origin of samples; the letter after the species name is the abbreviation used for geographical origin. The map is modified from ref. 30. **b**, Maximum-likelihood trees based on all autosomal sites; all nodes having full

local support on the basis of the Shimodaira–Hasegawa test are marked by asterisks. The colour code for groups of species applies to both panels. Taxa that showed deviations from classical taxonomy are underscored.

an earlier version of the taxonomy, in which these three groups were classified as distinct species on the basis of morphological differences^{17,18}.

Second, *Geospiza conirostris* on Española showed the highest genetic similarity to another species, *Geospiza magnirostris*, whereas *G. conirostris* on Genovesa clustered with *Geospiza scandens* (Fig. 1b). Here, phenotypic similarity parallels genetic similarity; *G. conirostris* on Genovesa have a pointed beak similar to *G. scandens*, whereas those on Española have a blunt beak more similar to the beaks of *G. magnirostris* (Extended Data Fig. 3).

A network constructed from autosomal genome sequences indicates conflicting signals in the internal branches of ground and tree finches that may reflect incomplete lineage sorting and/or gene flow (Extended Data Fig. 3). The exact branching order of the most recently evolved ground and tree finches should be interpreted with caution as it may change with additional sampling. Since our data revealed some important discrepancies with the phenotype-based taxonomy, we propose a revised taxonomy for the sharp-beaked ground finch (*G. difficilis*) and the large cactus finch (*G. conirostris*) (Supplementary Text and Extended Data Fig. 4), but will use the current names in the text.

We dated phylogenetic splits on the basis of genome divergence (Fig. 2a), and compared these estimates with those obtained using mtDNA (Extended Data Fig. 5a and Supplementary Text). We infer that the most basal split, between warbler finches (*Certhidea* sp.) and other finches, occurred about 900,000 years ago. The rapid radiations of ground and tree finches began around 100,000–300,000 years ago. Although these estimates are based on whole-genome data, they should be considered minimum times, as they do not take into account gene flow.

Extensive interspecies gene flow

The discrepancies between phylogenies based on morphology and genome sequences may be due to convergent evolution and/or interspecies gene flow. We found evidence of introgression from three sources: ABBA–BABA tests, discrepancies between phylogenetic trees based on autosomal and sex-linked loci, and mtDNA (Supplementary Text and Extended Data Fig. 5a).

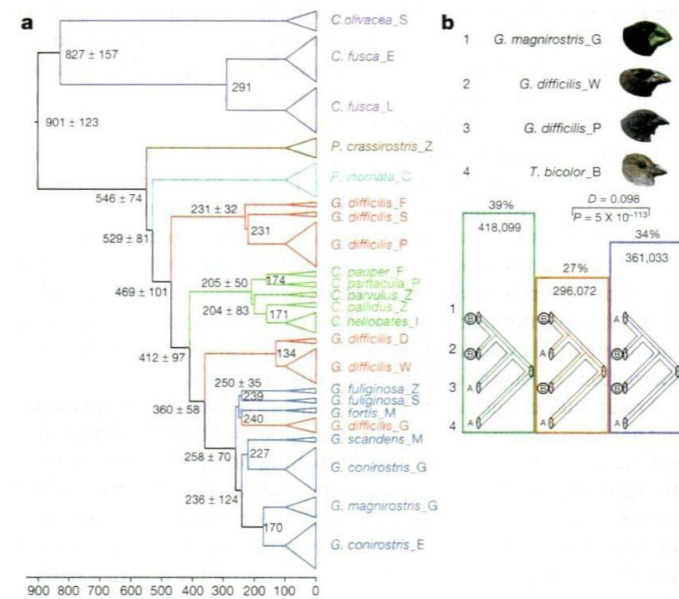


Figure 2 | Population history. **a**, Dating the nodes (in thousands of years) with confidence intervals (when applicable) in the phylogeny on the basis of divergence corrected for coalescence in ancestral populations; the topology is the representation of the inferred species tree from Fig. 1b. **b**, ABBA–BABA analysis of *G. magnirostris*, *G. difficilis* from Wolf and Pinta, and *L. noctis*. Number of sites supporting different trees is indicated both as a percentage and as actual numbers. The *D* statistic and corresponding Holm–Bonferroni-corrected *P* value are given for testing the null hypothesis of symmetry in genetic relationships. Finch heads are reproduced from ref. 5. *How and Why Species Multiply: The Radiation of Darwin's Finches* by Peter R. Grant & B. Rosemary Grant. Copyright © 2008 Princeton University Press. Reprinted by permission.

First, the D statistic¹⁹ associated with the ABBA–BABA test was used to compare two populations of *G. difficilis* from Pinta and Wolf, and *G. magnirostris* from Genovesa, using *L. noctis* as outgroup; *G. magnirostris* also occurs on Wolf but we lacked samples from that population. The analysis confirmed that *G. difficilis* on Wolf has a closer genetic relationship with *G. magnirostris* than with *G. difficilis* on Pinta (Fig. 2b). But there is evidence of gene flow between *G. difficilis* on Wolf and Pinta ($P = 5 \times 10^{-113}$), because the substantial asymmetry in genetic relationships cannot be explained by incomplete lineage sorting. However, the D statistic does not distinguish admixture from ancestral subdivision¹⁹. We conclude that the closely related populations of *G. difficilis* on Wolf and Darwin are a species of mixed ancestry where most of the genome originates from *G. magnirostris* or a close relative (Supplementary Table 2), whereas a considerable proportion of the genome, possibly including genetic variants affecting phenotypic characters, is derived from *G. difficilis*. Similarly, *G. difficilis* on Genovesa shows a closer genetic relationship to the other ground and tree finches than to *G. difficilis* on Pinta, but we also found evidence for gene flow between the two groups previously classified as *G. difficilis* ($P = 3 \times 10^{-87}$; Supplementary Table 2).

We next investigated gene flow involving the populations of *G. conirostris* on Genovesa and Española, which appear as separate species in our phylogenetic analysis. The ABBA–BABA analysis confirmed that *G. conirostris* on Española shows a closer genetic relationship to *G. magnirostris* than to *G. conirostris* on Genovesa (Extended Data Fig. 6a), but also provided evidence for gene flow between *G. conirostris* on Española and *G. conirostris* on Genovesa, which may explain some of their phenotypic similarities and their previous classification as a single species.

Given the evidence of relatively recent hybridization, we explored the possibility of more ancient hybridization between warbler finches (*Certhidea fusca* and *Certhidea olivacea*) and other finches. ABBA–BABA analysis provided evidence for gene flow between *C. fusca* and the other finches ($P = 7 \times 10^{-199}$; Extended Data Fig. 6b). This pattern of gene flow was apparent for all non-warbler finches, implying that it occurred before the radiation of the non-warbler finches (Supplementary Table 2).

The trees based on autosomal (Fig. 1b) and Z-linked sites (Extended Data Fig. 5b) are not completely congruent. The tree based on Z-linked polymorphisms indicated that *G. difficilis* present on the highlands of Pinta, Fernandina and Santiago is more closely related to *Platyspiza crassirostris* and emerged before the Cocos finch split off from the ground and tree finches, whereas the autosomal tree indicates a reversed order for the emergence of the two species. This discrepancy can potentially be explained by gene flow between *G. difficilis* and tree and ground finches after the Cocos finch became reproductively isolated from the finches on the Galápagos, which affected Z-linked and autosomal loci to different degrees. It is a common observation in closely related species that there is more interspecies sharing of sequence polymorphisms at autosomal loci than at sex-linked loci²⁰. This interpretation of the phylogenetic status of *G. difficilis* (highland group) is supported by the trees based on both mtDNA and W (Extended Data Fig. 5), which suggest that *G. difficilis* diverged from the ancestor of other ground and tree finches before the emergence of the Cocos finch.

Finally, our analysis of demographic history using the pairwise sequentially Markovian coalescent (PSMC) model²¹ was consistent with extensive interspecies gene flow among the ground finches, as they have maintained larger effective population sizes than the other species (Supplementary Text and Extended Data Fig. 6c, d).

A major locus controlling beak shape

The most striking morphological difference among Darwin's finches concerns beak shape (Extended Data Fig. 3). We performed a genome-wide scan on the basis of populations that are closely related but show different beak morphology: *G. magnirostris* and *G. conirostris* on Española have blunt beaks, whereas *G. conirostris* on Genovesa and *G. difficilis*

on Wolf have pointed beaks. We used non-overlapping 15-kilobase (kb) windows to identify regions with the highest fixation indices (F_{ST}) between groups. The F_{ST} distribution was Z-transformed (ZF_{ST}) and regions with striking ZF_{ST} values were identified (Fig. 3a). Among the 15 most significant regions, six harboured genes previously associated with craniofacial and/or beak development in mammals or birds including calmodulin (*CALM*)¹¹, goosecoid homeobox (*GSC*)²², retinol dehydrogenase 14 (*RDH14*)²³, ALX homeobox 1 (*ALX1*)^{24,25}, fibroblast growth factor 10 (*FGF10*)²⁶ and forkhead box C1 (*FOXO1*)²⁷. A previous study demonstrated differential expression of *CALM* between finches with different beak types¹¹. Two other studies reported differential expression of bone morphogenetic protein 4 (*BMP4*)^{10,12}, but we did not observe any elevated ZF_{ST} values in the vicinity of this locus, suggesting that differential expression is controlled by other loci.

The most striking finding was a 240-kb region with high ZF_{ST} values, including the window with the highest ZF_{ST} score (9.46) overall (Fig. 3a, b). The region overlaps part of *LRR1Q1* (leucine-rich repeats and IQ motif containing 1), the entire *ALX1* gene and about 130 kb downstream of *ALX1*. No previous report indicates that *LRR1Q1* has a role during development in vertebrates. By contrast, *ALX1* is an excellent candidate for variation in beak morphology. It encodes a paired-type homeodomain protein that plays a crucial role in development of structures derived from craniofacial mesenchyme, the first branchial arch and the limb bud²⁴, and on migration of cranial neural crest cells, highly relevant to beak development²⁵. Loss of *ALX1* in humans causes disruption of early craniofacial development²⁴.

All individuals in the blunt beak category were homozygous for a blunt beak-associated haplotype (denoted *B*), except one heterozygous *G. conirostris* individual from Española. Furthermore, except for one heterozygous bird from Genovesa, all 19 *G. difficilis* individuals not included in the F_{ST} scan were homozygous for a pointed beak haplotype (*P*), consistent with their phenotypic appearance (sharp-beaked ground finches). This is notable because genome-wide, *G. difficilis* on Wolf, Darwin and Genovesa are all more closely related to the blunt-beaked *G. magnirostris* than to the pointed-beaked *G. difficilis* from Pinta (Fig. 2b).

A phylogenetic tree based on this region revealed a deep divergence between the *B* and *P* haplotypes that must have occurred soon after the split between warbler finches and other Darwin's finches (Fig. 3c). Apart from the blunt-beaked *G. magnirostris* and *G. conirostris* on Española, all individuals except three were homozygous for *P* haplotypes, the remaining three being heterozygous. The two *G. fortis* from Daphne Major Island were both homozygous, but for different haplotypes (*BB* and *PP*; Fig. 3c). The short branch lengths among *B* haplotypes are consistent with a selective sweep. There were 335 fixed differences between the *B* and *P* haplotypes (Fig. 3d, upper panel), which we assigned as derived or ancestral on the basis of comparison with the outgroup sequence (*L. noctis*). Derived alleles on the *B* haplotype were aggregated in the vicinity of *ALX1*, including the downstream region (Fig. 3d, middle panel). Furthermore, 8 of these 335 fixed differences occurred at conserved sites, and the *B* haplotype carried the derived allele at seven of them (Fig. 3d, lower panel). Four derived alleles occurred at sites corresponding to transcription factor binding sites in the human genome²⁸. Two other changes constitute missense mutations (L112P and I208V) at *ALX1* amino-acid residues that are highly conserved among birds and mammals (Extended Data Fig. 7), and 'Sorting Intolerant From Tolerant' (SIFT)²⁹ analysis classified both as damaging (score 0.03 for both). The ratio of non-synonymous to synonymous substitutions between the *P* and *B* alleles is high (2/1 = 2.00) compared with the ratio observed between the ancestral *P* allele and orthologous zebra finch (2/14 = 0.14) and human (21/122 = 0.17) sequences, suggesting that one or both of these missense mutations are non-neutral.

That *ALX1* is polymorphic in *G. fortis* (Fig. 3c, d, upper panel) is particularly interesting, because field observations have shown there is considerable diversity in beak shape in this species^{5,30}. We genotyped an additional 62 *G. fortis* birds from Daphne Major Island for a diagnostic

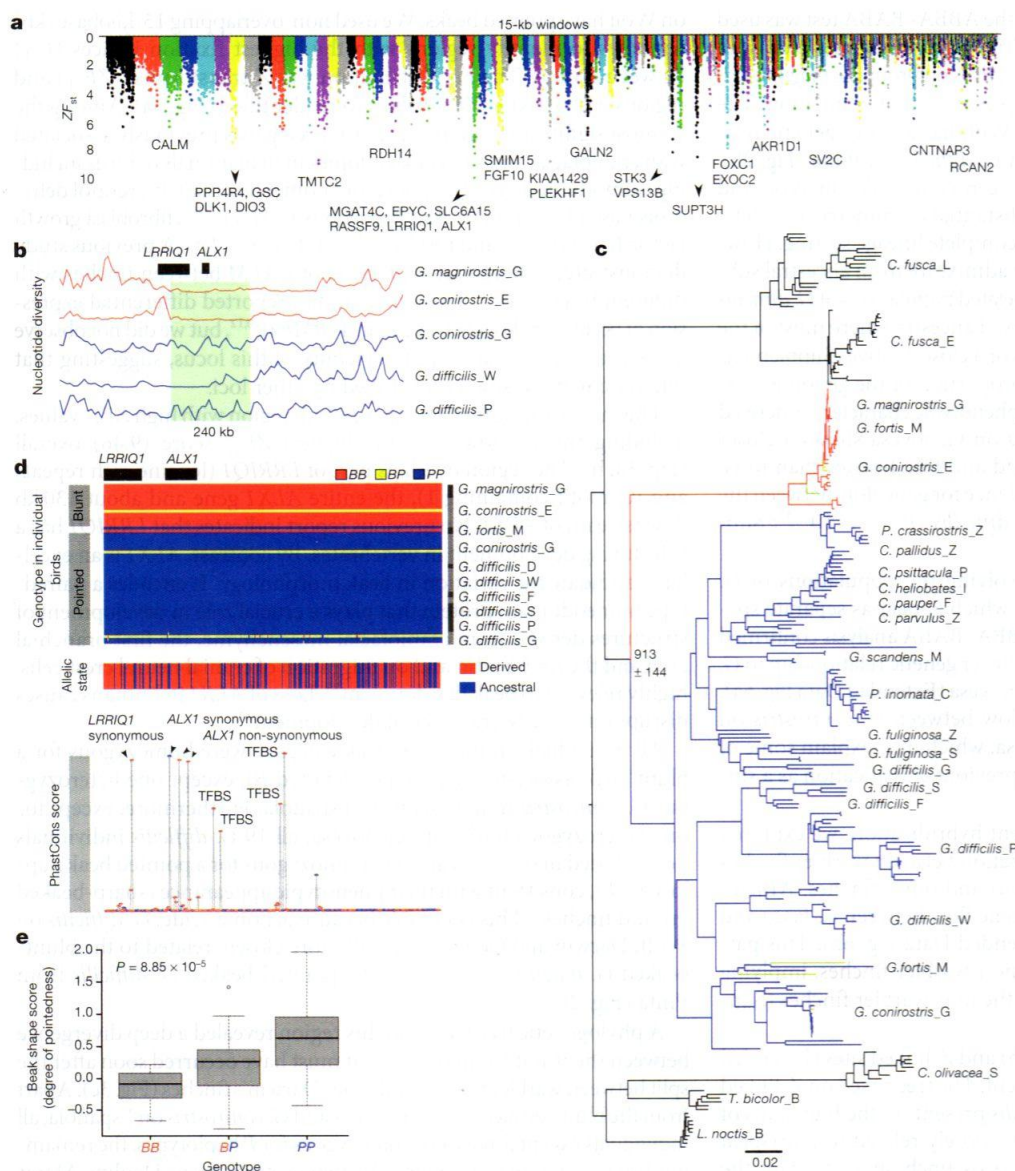


Figure 3 | A major locus controlling beak shape. **a**, Genome-wide F_{ST} screen comparing *G. magnirostris* and *G. conirostris* (Española) having blunt beaks with *G. conirostris* (Genovesa) and *G. difficilis* (Wolf) having pointed beaks. The y axis represents ZF_{ST} values. **b**, Nucleotide diversities in the *ALX1* region. The 240-kb region showing high homozygosity in blunt-beaked species is highlighted. Red and blue colours in **b–d** refer to blunt and pointed beak haplotypes, respectively. **c**, Neighbour-joining haplotype tree of *ALX1* region. Haplotypes originating from heterozygous birds (see text) are indicated in yellow. Estimated time since divergence (\pm confidence interval) of blunt and pointed beak haplotypes are given in thousands of years. **d**, Upper panel: genotypes at 335 SNPs showing complete fixation between *ALX1* haplotypes associated with blunt (B) and pointed (P) beaks. **d**, Middle panel: classification of alleles associated with blunt beaks at the 335 SNPs as derived or ancestral on the basis of allelic state in the outgroup. **d**, Lower panel: PhastCons³⁵ scores (on the basis of human, mouse and finch alignments) for the 335 SNP sites. TFBS, transcription factor binding sites. **e**, Linear regression analysis of beak-shape scores among *G. fortis* individuals on Daphne Major Island classified according to *ALX1* genotype; distribution of pointedness in each class is shown as a boxplot; $n = 62$; $F = 17.7$, adjusted $R^2 = 0.22$. Differences in six individual body and beak size traits were not significant (all $P > 0.05$).

single nucleotide polymorphism (SNP), and observed a significant association with beak shape ($P = 8.8 \times 10^{-5}$, Fig. 3e). *PP* homozygotes tended to have proportionately long, pointed beaks, *BB* homozygotes had proportionately deep, blunt beaks, whereas heterozygotes (*BP*) had intermediate beak shapes. We also compared haplotype frequencies among *G. fortis* individuals on Daphne Major Island with those on Santa Cruz, which have a larger and blunter beak on average³¹, possibly as a result of introgressive hybridization with *G. magnirostris*^{4,5}. We found the *B* haplotype to be more frequent on Santa Cruz than on Daphne Major (0.74, $n = 21$ versus 0.49, $n = 62$; $P = 0.007$, Fisher's exact test).

Natural selection on beak size and shape of *G. fortis* on Daphne Major Island has led to evolutionary change in the past few decades^{5,30}. Moreover, genetic variation in beak shape has been increased through introgressive hybridization^{5,30} with two species of *Geospiza*, *scandens* and *fuliginosa*, that have relatively pointed beaks. Therefore we expect hybrids and backcrosses in the *G. fortis* population to have a relatively high frequency of the *P* haplotype. We genotyped an additional 25 *G. fortis* at *ALX1*, added them to the sample of 62 (Methods) and compared the haplotype frequencies in eight hybrids (including backcrosses) and 79 non-hybrids. *ALX1-P* had a frequency of 0.75 among hybrids, and 0.44 among the others, which is statistically significant in the expected direction ($P = 0.03$, Fisher's exact test). Thus, *ALX1-P* alleles

introduced by introgressive hybridization most probably contributed to evolution of more pointed beaks in 1987 following natural selection as a result of a change in food supply in the 1985–86 drought³⁰.

Discussion

Our revised and dated phylogeny of Darwin's finches shows that the adaptive radiation took place in the past million years, with a rapid accumulation of species recently (Supplementary Text). We have genomically characterized the entire radiation, which has revealed a striking connection between past and present evolution. Evidence of introgressive hybridization, which has been documented as a contemporary process, is found throughout the radiation. Hybridization has given rise to species of mixed ancestry, in the past (this study) and the present³⁰. It has influenced the evolution of a key phenotypic trait: beak shape. Similar introgressive hybridization affecting an adaptive trait (mimicry) has been described in *Heliconius* butterflies³². The degree of continuity between historical and contemporary evolution is unexpected because introgressive hybridization plays no part in traditional accounts of adaptive radiations of animals^{1,2}. For young radiations it complements the better-known role of natural selection.

Charles Darwin first noted the diversity in beak shapes among the finches on Galápagos. Our genomic study has now revealed some of

the underlying genetic variation explaining this diversity. A polygenic basis for beak diversity is indicated by our discovery of about 15 regions with strong genetic differentiation between groups of finches with blunt or pointed beaks. We present evidence that the *ALX1* locus contributes to beak diversity, within and among species. The derived *ALX1-B* haplotype associated with blunt beaks has a long evolutionary history (hundreds of thousands of years), because its origin predates the radiation of vegetarian, tree and ground finches (Fig. 3c). This haplotype is fixed or nearly fixed in two ground finches with blunt beaks, *G. magnirostris* and *G. conirostris* on Española, and it co-segregates with variation in beak shape in *G. fortis*. As previously documented in domestic animals³³ and natural populations³⁴, the haplotype might have evolved by accumulating both coding and regulatory changes affecting *ALX1* function. Natural selection and introgression affecting this locus have contributed to the diversification of beak shapes among Darwin's finches and hence to their expanded utilization of food resources on Galápagos.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 October; accepted 31 December 2014.

Published online 11 February 2015.

- Schluter, D. *The Ecology of Adaptive Radiation* (Oxford Univ. Press, 2000).
- Seehausen, O. African cichlid fish: a model system in adaptive radiation research. *Proc. R. Soc. B* **273**, 1987–1998 (2006).
- Lack, D. *Darwin's Finches* (Cambridge Univ. Press, 1947).
- Grant, P. R. *Ecology and Evolution of Darwin's Finches* (Princeton Univ. Press, 1999).
- Grant, P. R. & Grant, B. R. *How and Why Species Multiply. The Radiation of Darwin's Finches* (Princeton Univ. Press, 2008).
- Petren, K., Grant, P. R., Grant, B. R. & Keller, L. F. Comparative landscape genetics and the adaptive radiation of Darwin's finches: the role of peripheral isolation. *Mol. Ecol.* **14**, 2943–2957 (2005).
- Ali, J. R. & Aitchison, J. C. Exploring the combined role of eustasy and oceanic island thermal subsidence in shaping biodiversity on the Galápagos. *J. Biogeogr.* **41**, 1227–1241 (2014).
- Geist, D., Snell, H., Snell, H., Goddard, C. & Kurz, M. In *The Galápagos: A Natural Laboratory for the Earth Sciences* (eds Harpp K. S., Mittelstaedt E., d'Ozouville N., & Graham, D.) 145–166 (American Geophysical Union, 2014).
- Farrington, H. L., Lawson, L. P., Clark, C. M. & Petren, K. The evolutionary history of Darwin's finches: speciation, gene flow, and introgression in a fragmented landscape. *Evolution* **68**, 2932–2944 (2014).
- Abzhanov, A., Protas, M., Grant, B. R., Grant, P. R. & Tabin, C. J. Bmp4 and morphological variation of beaks in Darwin's finches. *Science* **305**, 1462–1465 (2004).
- Abzhanov, A. *et al.* The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* **442**, 563–567 (2006).
- Mallarino, R. *et al.* Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proc. Natl Acad. Sci. USA* **108**, 4057–4062 (2011).
- Burns, K. J. *et al.* Phylogenetics and diversification of tanagers (Passeriformes: Thraupidae), the largest radiation of Neotropical songbirds. *Mol. Phylogenet. Evol.* **75**, 41–77 (2014).
- Zhang, G., Parker, P., Li, B., Li, H. & Wang, J. The genome of Darwin's finch (*Geospiza fortis*). *GigaScience*, <http://dx.doi.org/10.5524/100040> (3 August 2012).
- Ellegren, H. The evolutionary genomics of birds. *Annu. Rev. Ecol. Syst.* **44**, 239–259 (2013).
- Balakrishnan, C. N. & Edwards, S. V. Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* **181**, 645–660 (2009).
- Swarth, H. S. The avifauna of the Galapagos Islands. *Occ. Pap. Calif. Acad. Sci.* **18**, 1–299 (1931).
- Lack, D. The Galapagos finches (Geospizinae): a study in variation. *Occ. Pap. Calif. Acad. Sci.* **21**, 1–159 (1945).
- Durand, E. Y., Patterson, N., Reich, D. & Slatkin, M. Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* **28**, 2239–2252 (2011).
- Qvarnstrom, A. & Bailey, R. I. Speciation through evolution of sex-linked genes. *Heredity* **102**, 4–15 (2009).
- Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).
- Rivera-Perez, J. A., Wakamiya, M. & Behringer, R. R. Goosecoid acts cell autonomously in mesenchyme-derived tissues during craniofacial development. *Development* **126**, 3811–3821 (1999).
- Rowe, A., Richman, J. M. & Brickell, P. M. Retinoic acid treatment alters the distribution of retinoic acid receptor- β transcripts in the embryonic chick face. *Development* **111**, 1007–1016 (1991).
- Uz, E. *et al.* Disruption of *ALX1* causes extreme microphthalmia and severe facial clefting: expanding the spectrum of autosomal-recessive *ALX*-related frontonasal dysplasia. *Am. J. Hum. Genet.* **86**, 789–796 (2010).
- Dee, C. T., Szymoniuk, C. R., Mills, P. E. D. & Takahashi, T. Defective neural crest migration revealed by a zebrafish model of *Alx1*-related frontonasal dysplasia. *Hum. Mol. Genet.* **22**, 239–251 (2013).
- Brugmann, S. A. *et al.* Comparative gene expression analysis of avian embryonic facial structures reveals new candidates for human craniofacial disorders. *Hum. Mol. Genet.* **19**, 920–930 (2010).
- Sommer, P., Napier, H. R., Hogan, B. L. & Kidson, S. H. Identification of *Tgfb1i4* as a downstream target of *Foxc1*. *Dev. Growth Differ.* **48**, 297–308 (2006).
- Wang, J. *et al.* Factorbook.org: a Wiki-based database for transcription factor-binding data generated by the ENCODE consortium. *Nucleic Acids Res.* **41**, D171–D176 (2013).
- Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols* **4**, 1073–1081 (2009).
- Grant, P. R. & Grant, B. R. *40 Years of Evolution. Darwin's Finches on Daphne Major Island* (Princeton Univ. Press, 2014).
- Boag, P. T. Growth and allometry of external morphology in Darwin's finches (*Geospiza*) on Isla Daphne Major, Galápagos. *J. Zool.* **204**, 413–441 (1984).
- The Heliconius Genome Consortium. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98 (2012).
- Andersson, L. Molecular consequences of animal breeding. *Curr. Opin. Genet. Dev.* **23**, 295–301 (2013).
- Linnen, C. R. *et al.* Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339**, 1312–1316 (2013).
- Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements The National Science Foundation (USA) funded the collection of material under permits from the Galápagos and Costa Rica National Parks Services, and in accordance with protocols of Princeton University's Animal Welfare Committee. The map and images of finch heads are reproduced with permission from Princeton University Press. The project was supported by the Knut and Alice Wallenberg Foundation. Sequencing was performed by the SNP&SEQ Technology Platform, supported by Uppsala University and Hospital, SciLifeLab and Swedish Research Council (80576801 and 70374401). Computer resources were supplied by UPPMAX.

Author Contributions P.R.G. and B.R.G. collected the material. L.A., P.R.G. and B.R.G. conceived the study. L.A. and M.T.W. led the bioinformatic analysis of data. S.L. and J.B. performed the bioinformatic analysis with contributions from M.S.A., K.M., M.G., A.M.-B., C.-J.R. and N.Z. M.P. and C.W. performed experimental work. L.A., S.L., J.B., B.R.G., P.R.G. and M.T.W. wrote the paper with input from the other authors. All authors approved the manuscript before submission.

Author Information The Illumina reads have been submitted to the short reads archive (<http://www.ncbi.nlm.nih.gov/sra>) under accession number PRJNA263122 and the consensus sequence for the *G. fortis* mtDNA has been submitted to GenBank under accession number KM891730. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.A. (leif.andersson@imbim.uu.se).