Association with Microcephaly and Guillain–Barré Syndrome. Third Update, 23 February 2016 (ECDC, 2016).
8. B. D. Lindenbach, C. L. Murray, H.-J. Thile, C. M. Rice, in *Fields Virology*, D. M. Knipe, P. M. Howley, Eds. (Lippincott Williams & Wilkins, ed. 6, vol. 1, 2013), chap. 25, pp. 712–746.
9. J. Mlakar *et al.*, *N. Engl. J. Med.* **374**, 951–958 (2016).
10. R. B. Martines *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **65**, 159–160 (2016).
11. V.-M. Cao-Lormeau *et al.*, *Lancet* 10.1016/S0140-6736(16)00562-6 (2016).
12. B. D. Foy *et al.*, *Emerg. Infect. Dis.* **17**, 880–882 (2011).
13. D. Musso *et al.*, *Emerg. Infect. Dis.* **21**, 359–361 (2015).
14. Y. Zhang *et al.*, *EMBO J.* **22**, 2604–2613 (2003).
15. R. J. Kuhn *et al.*, *Cell* **108**, 717–725 (2002).
16. X. Zhang *et al.*, *Nat. Struct. Mol. Biol.* **20**, 105–110 (2013).
17. I. M. Yu *et al.*, *Science* **319**, 1834–1837 (2008).
18. X. Zhang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6795–6799 (2013).
19. T. C. Pierson, M. S. Diamond, *Curr. Opin. Virol.* **2**, 168–175 (2012).
20. C. Baronti *et al.*, *Genome Announc.* **2**, e00500-14 (2014).
21. A. Enfissi, J. Codrington, J. Roosblad, M. Kazanji, D. Rousset, *Lancet* **387**, 227–228 (2016).
22. R. S. Lanciotti *et al.*, *Emerg. Infect. Dis.* **14**, 1232–1239 (2008).
23. G. C. Lander *et al.*, *J. Struct. Biol.* **166**, 95–102 (2009).
24. S. H. Scheres, *J. Struct. Biol.* **180**, 519–530 (2012).
25. P. B. Rosenthal, R. Henderson, *J. Mol. Biol.* **333**, 721–745 (2003).
26. F. Guo, W. Jiang, *Methods Mol. Biol.* **1117**, 401–443 (2013).
27. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
28. P. V. Afonine *et al.*, *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
29. A. T. Brünger *et al.*, *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
30. S. Mukhopadhyay, B. S. Kim, P. R. Chipman, M. G. Rossmann, R. J. Kuhn, *Science* **302**, 248 (2003).
31. E. Pokidysheva *et al.*, *Cell* **124**, 485–493 (2006).
32. J. L. Miller *et al.*, *PLOS Pathog.* **4**, e17 (2008).
33. S. M. Lok, *Trends Microbiol.* 10.1016/j.tim.2015.12.004 (2015).
34. D. W. Beasley *et al.*, *J. Virol.* **79**, 8339–8347 (2005).
35. O. Faye *et al.*, *PLOS Negl. Trop. Dis.* **8**, e2636 (2014).
36. S. Nelson *et al.*, *PLOS Pathog.* **4**, e1000060 (2008).

ACKNOWLEDGMENTS

We thank X. de Lamballerie (Emergence des Pathologies Virales, Aix-Marseille Université, Marseille, France) and the European Virus Archive Goes Global (EVAg) for consenting to the use of the H/PF/2013 ZIKV strain for this study under a material transfer agreement with EVAg's partner, Aix-Marseille Université, and we thank M. S. Diamond (Washington University, St. Louis, MO, USA) for sending the virus. We acknowledge E. Frye for help with sequence alignments. We also acknowledge the use of the Purdue cryo-EM facility. We are grateful to W. Jiang for his jspr program, which was used to perform the anisotropic magnification corrections, and Y. Liu for providing help in submitting the Protein Data Bank (PDB) coordinates. The work presented in this report was funded by the National Institute of Allergy and Infectious Diseases of the NIH through grants R01AI073755 and R01AI076331 to M.G.R. and R.J.K. T.C.P was supported by the intramural program of the National Institute of Allergy and Infectious Diseases. Supporting information for this research is provided in the supplementary materials. The atomic coordinates and cryo-EM density maps for the mature ZIKV are available at the PDB and Electron Microscopy Data Bank under accession codes 5IRE and EMD-8116, respectively.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6284/467/suppl/DC1
Figs. S1 and S2
Tables S1 and S2
References (*37*, *38*)

4 March 2016; accepted 21 March 2016
Published online 31 March 2016
10.1126/science.aaf5316

EVOLUTIONARY GENETICS

A beak size locus in Darwin's finches facilitated character displacement during a drought

Sangeet Lamichhaney,[1] Fan Han,[1] Jonas Berglund,[1] Chao Wang,[1] Markus Sällman Almén,[1] Matthew T. Webster,[1] B. Rosemary Grant,[2] Peter R. Grant,[2] Leif Andersson[1,3,4]*

Ecological character displacement is a process of morphological divergence that reduces competition for limited resources. We used genomic analysis to investigate the genetic basis of a documented character displacement event in Darwin's finches on Daphne Major in the Galápagos Islands: The medium ground finch diverged from its competitor, the large ground finch, during a severe drought. We discovered a genomic region containing the *HMGA2* gene that varies systematically among Darwin's finch species with different beak sizes. Two haplotypes that diverged early in the radiation were involved in the character displacement event: Genotypes associated with large beak size were at a strong selective disadvantage in medium ground finches (selection coefficient $s = 0.59$). Thus, a major locus has apparently facilitated a rapid ecological diversification in the adaptive radiation of Darwin's finches.

Similar species potentially compete for limited resources when they encounter each other through a change in geographical ranges. As a result of resource competition, they may diverge in traits associated with exploiting these resources (*1*, *2*). Darwin proposed this as the principle of character divergence [now known as ecological character displacement (*3*, *4*)], a process invoked as an important mechanism in the assembly of complex ecological communities (*5*, *6*). It is also an important component of models of speciation (*6*, *7*). However, it has been difficult to obtain unequivocal evidence for ecological character displacement in nature (*8*, *9*). The medium ground finch (*Geospiza fortis*) and large ground finch (*G. magnirostris*) on the small island of Daphne Major provide one example where rigorous criteria have been met (*10*). Beak sizes diverged as a result of a selective disadvantage to medium ground finches with large beaks when food availability declined through competition with large ground finches during a severe drought in 2004–2005 (*11*).

Size-related traits can pose problems for the analysis of selection, and Darwin's finch beaks are no exception, as beak size and body size are strongly correlated ($r = 0.7$ to 0.8) (*11*). We used a combination of multiple regression and selection differential analysis to investigate the 2004–2005 selection event. Statistically, these produced much stronger associations between survival and beak size ($S = -1.02$, $P < 0.0001$) than between survival and body size ($S = -0.67$, $P < 0.05$). Thus, body size was possibly subject to selection, but beak size was a more important factor affecting the probability of survival independent of body size (*11*, *12*). However, the genetic basis of the selected traits remains unknown. Beak dimensions and overall body size of the medium ground finch are highly heritable (*13*), but no gene(s) regulating body size have been identified. Furthermore, although some signaling molecules affecting beak dimensions in Darwin's finches have been identified (*14*), only one regulatory gene, *ALX1*, is known and it regulates variation in beak shape (*15*), which was not associated with survival in 2004–2005.

We performed a genome-wide screen for loci affecting overall body size in six species of Darwin's finches that primarily differ in size and size-related traits: the small, medium, and large ground finches, and the small, medium, and large tree finches (Fig. 1, A and B, and table S1). Ground finches and tree finches diverged about 400,000 years ago and exhibit ongoing gene flow within and between the two groups (*15*). By combining species of similar size in different taxa, we minimized phylogenetic effects when contrasting the genomes of species differing in size. We then genotyped individuals of the Daphne population of medium ground finches that succumbed or survived during the drought of 2004–2005. This approach allowed us to identify a locus with major effect on beak size variation that played a key role in the character displacement episode.

We sequenced 10 birds from each of the six species (total 60 birds) to ~10× coverage per individual, using 2 × 125–base pair paired-end reads. The sequences were aligned to the reference genome from a female medium ground

[1]Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden. [2]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. [3]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden. [4]Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA.
***Corresponding author. Email: leif.andersson@imbim.uu.se**

Downloaded from http://science.sciencemag.org/ on April 22, 2016

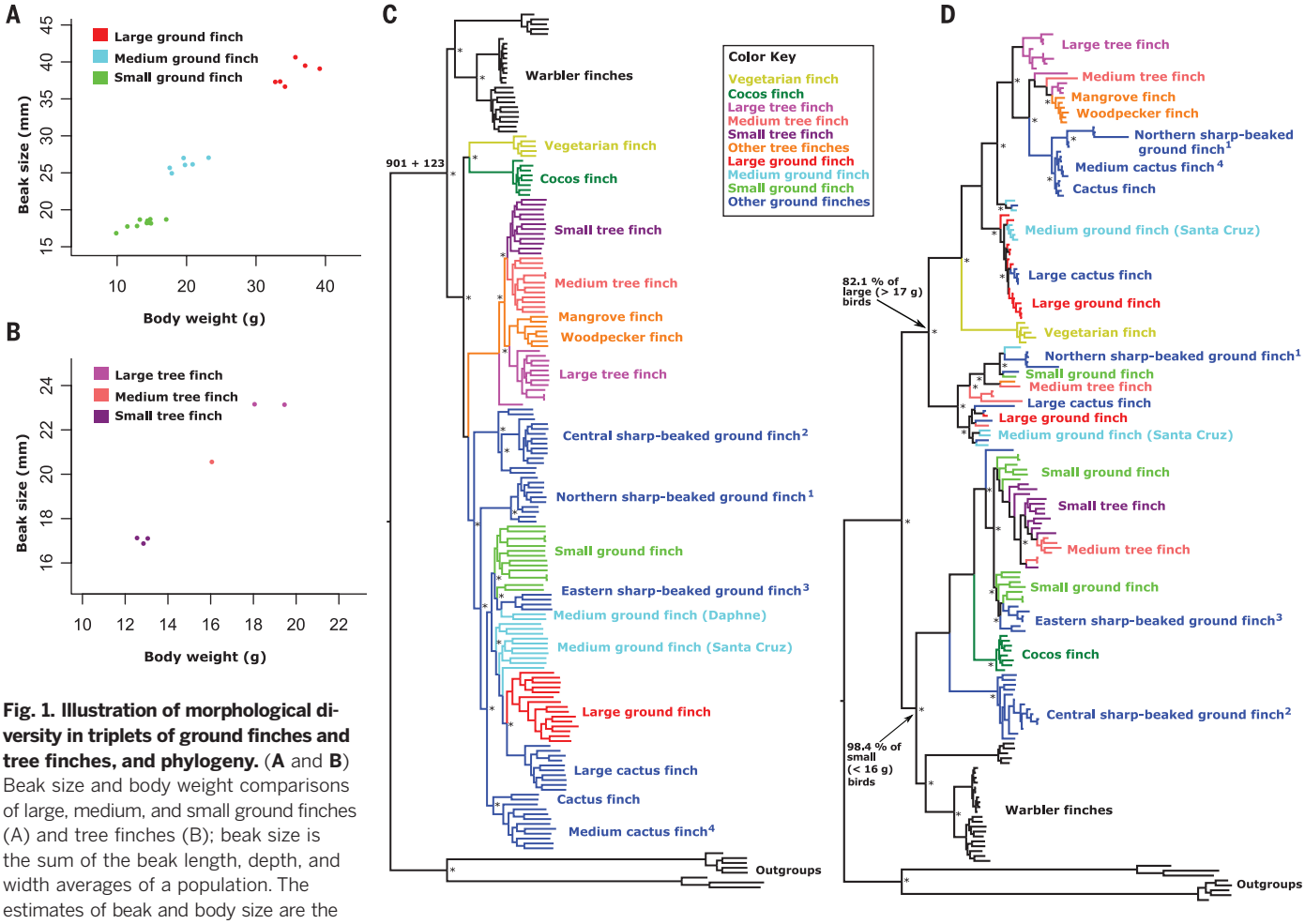


Fig. 1. Illustration of morphological diversity in triplets of ground finches and tree finches, and phylogeny. (**A** and **B**) Beak size and body weight comparisons of large, medium, and small ground finches (A) and tree finches (B); beak size is the sum of the beak length, depth, and width averages of a population. The estimates of beak and body size are the population averages; each dot represents a population from a specific island in the Galápagos archipelago. (**C**) Maximum-likelihood tree using all polymorphic autosomal sites. The estimated divergence time, with its 95% confidence interval based on nuclear sites (*15*) between warbler and nonwarbler finches, is shown in thousands of years; the corresponding estimate for this split on the basis of mitochondrial DNA cytochrome b sequences is 1.4 ± 0.2 million years (*15*). (**D**) Maximum-likelihood tree from the 525-kb region around *HMGA2*. All nodes having full local support on the basis of the Shimodaira-Hasegawa test are marked by asterisks. For sharp-beaked ground finches and medium cactus finch from Genovesa, the revised taxonomy as proposed in (*15*) is used: [1]northern sharp-beaked ground finch from Wolf and Darwin (*Geospiza septentrionalis*); [2]central sharp-beaked ground finch from Pinta, Santiago, and Fernandina (*G. difficilis*); [3]eastern sharp-beaked ground finch from Genovesa (*G. acutirostris*); [4]medium cactus finch from Genovesa (*G. propinqua*).

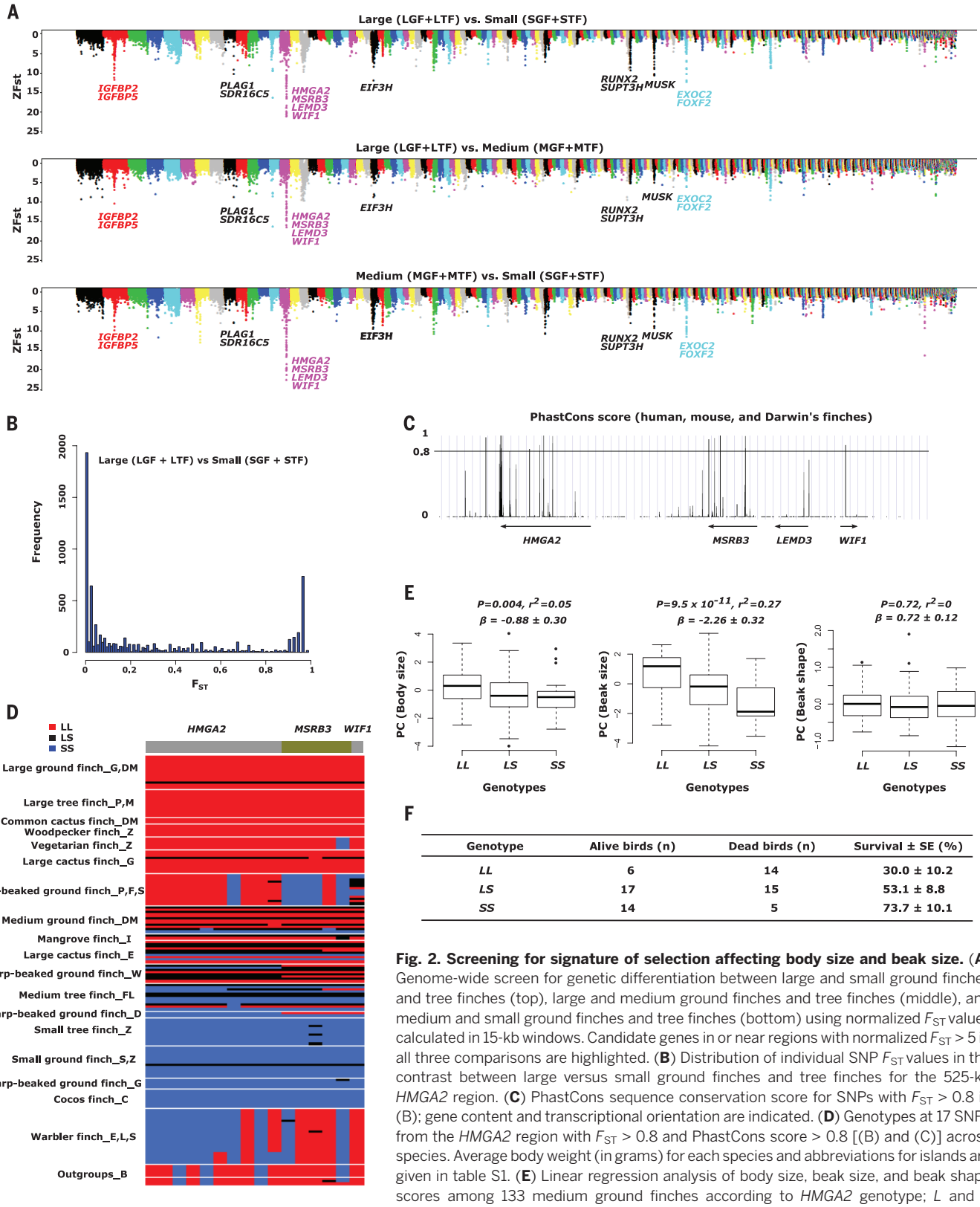
Table 1. Summary of 60 samples of large, medium, and small ground finches and tree finches used for whole-genome sequencing.

Common name	Species	No. of samples	Island	ID
Large ground finch	*Geospiza magnirostris*	10	Daphne	LGF
Medium ground finch	*Geospiza fortis*	10	Santa Cruz	MGF
Small ground finch	*Geospiza fuliginosa*	10	Santiago	SGF
Large tree finch	*Camarhynchus psittacula*	8	Pinta	LTF
		1	Marchena	
		1	Isabela	
Medium tree finch	*Camarhynchus pauper*	10	Floreana	MTF
Small tree finch	*Camarhynchus parvulus*	10	Santa Cruz	STF

finch (*12*). We combined these data with sequences from 120 birds, including all species of Darwin's finches and two outgroup species (*15*), to call 44,767,199 variable sites within or between populations after stringent variant calling. We constructed a maximum-likelihood phylogenetic tree on the basis of all 180 genome sequences (Fig. 1C). This tree was almost identical to our previous tree (*15*).

A genome-wide fixation index (F_{ST}) scan comparing large, medium, and small ground finches and tree finches (Table 1) identified seven independent genomic regions with consistent genetic differentiation ($ZF_{ST} > 5$) in each contrast (Fig. 2A and table S2). One of these regions (~525 kb in size) showed the strongest differentiation in all three contrasts. The region included four genes: *high mobility AT-hook 2* (*HMGA2*), *methionine sulfoxide reductase B3* (*MSRB3*), *LEM domain-containing protein 3* (*LEMD3*), and *WNT inhibitory factor 1* (*WIF1*). This signal was also detected in F_{ST} screens comparing large, medium, and small birds separately within ground and tree finches (fig. S1). HMGA2 is a chromatin-associated protein that appears to lack intrinsic transcriptional activity but potentiates the effect of other transcription factors (*16*). Because

Downloaded from http://science.sciencemag.org/ on April 22, 2016



F

Genotype	Alive birds (n)	Dead birds (n)	Survival ± SE (%)
LL	6	14	30.0 ± 10.2
LS	17	15	53.1 ± 8.8
SS	14	5	73.7 ± 10.1

Fig. 2. Screening for signature of selection affecting body size and beak size. (**A**) Genome-wide screen for genetic differentiation between large and small ground finches and tree finches (top), large and medium ground finches and tree finches (middle), and medium and small ground finches and tree finches (bottom) using normalized F_{ST} values calculated in 15-kb windows. Candidate genes in or near regions with normalized $F_{ST} > 5$ in all three comparisons are highlighted. (**B**) Distribution of individual SNP F_{ST} values in the contrast between large versus small ground finches and tree finches for the 525-kb *HMGA2* region. (**C**) PhastCons sequence conservation score for SNPs with $F_{ST} > 0.8$ in (B); gene content and transcriptional orientation are indicated. (**D**) Genotypes at 17 SNPs from the *HMGA2* region with $F_{ST} > 0.8$ and PhastCons score > 0.8 [(B) and (C)] across species. Average body weight (in grams) for each species and abbreviations for islands are given in table S1. (**E**) Linear regression analysis of body size, beak size, and beak shape scores among 133 medium ground finches according to *HMGA2* genotype; *L* and *S* represent alleles present in birds with large and small beaks, respectively. The distribution of respective morphometric scores in each genotype class is shown as a boxplot together with *P* values, r^2 scores, and linear regression slopes (β ± SE) from the regression analyses. (**F**) Survival (% ± SE) according to *HMGA2* genotype among 71 medium ground finches experiencing the severe drought in 2004–2005.

Downloaded from http://science.sciencemag.org/ on April 22, 2016

a loss-of-function mutation in *Hmga2* causes the pygmy phenotype in mice that exhibits severe growth retardation (*17*) and because *HMGA2* has been associated with variation in height, craniofacial distances, and primary tooth eruption in humans (*18*, *19*), *HMGA2* was identified as a candidate gene. We refer to this region as the *HMGA2* locus but note that it includes three additional genes that may contribute to phenotypic effects (*12*).

We constructed a maximum-likelihood phylogenetic tree on the basis of this ~525-kb region, which revealed two major haplotype groups associated with size; 98% of small birds (body weight <16 g) clustered into one group and 82% of the large birds (body weight >17 g) clustered into the other (Fig. 1D). The split between the two haplotypes occurred before the divergence of warbler and nonwarbler finches at the base of the phylogeny (Fig. 1D), about 1 million years ago (Fig. 1C).

We calculated F_{ST} values per SNP (single-nucleotide polymorphism) for all SNPs within the ~525-kb *HMGA2* region (Fig. 2B). There were 1327 SNPs with strong genetic differentiation ($F_{ST} > 0.8$) spread across the region, but only one of these was coding (a missense mutation in *MSRB3*), which implies that most or all mutations causing the association with phenotype are regulatory. We identified 17 SNPs showing high genetic divergence between large and small ground finches and tree finches ($F_{ST} > 0.8$) at nucleotide sites in highly conserved regions across birds and mammals (PhastCons score > 0.8) (Fig. 2C). Six of these 17 SNPs cluster at the 3′ end of *HMGA2*. A comparison with the outgroup species (*Loxigilla noctis* and *Tiaris bicolor*) shows that the haplotype present in small birds is associated with the derived allele at a majority of these 17 SNPs (13/17; $P = 0.05$, binomial test). Large birds were homozygous for haplotypes belonging to one group, whereas the majority of small birds were homozygous for haplotypes belonging to the other group (Fig. 2D). Segregation is mainly observed in species with intermediate size (medium ground and tree finches).

Large, medium, and small ground finches and tree finches differ markedly both in body and beak size (Fig. 1, A and B, and table S1). Hence, we investigated whether the *HMGA2* locus is primarily associated with variation in body size, beak size, or both. As this locus shows segregation (Fig. 2D) in medium ground finches—a species with considerable diversity in both body and beak size (*10*)—we genotyped an additional 133 individuals of this species for a haplotype diagnostic SNP (A/G) at nucleotide position 7,003,776 base pairs in scaffold JH739900, ~2.3 kb downstream of *HMGA2*. This SNP showed a highly significant association with beak size, a significant association with body size, and no association with beak shape among medium ground finches (Fig. 2E). The locus appears to have an additive effect on beak size, where heterozygotes show an intermediate phenotype relative to the two homozygous classes, and linear regression analysis explains as much as 27% of the phenotypic variance in this population.

Six other loci showed consistent associations with overall size, but the genetic differentiation was not as pronounced as for the *HMGA2* locus (Fig. 2A). Interestingly, *PLAG1* and *SUPT3H* have previously been associated with height in humans (www.ebi.ac.uk/gwas), and *IGFBP2* encodes a protein that binds insulin-like growth factor I and II in plasma (Fig. 2A). All six loci were segregating in medium ground finches, but none showed a significant association with beak size, body size, or beak shape variation (fig. S2B). The results suggest that the phenotypic effects of these loci are small relative to the effect of the *HMGA2* locus.

We genotyped a diagnostic SNP for the *HMGA2* locus in medium ground finches on Daphne Major that experienced the severe drought in 2004–2005 ($n = 71$; 37 survived and 34 died) (*11*). Differential mortality resulted in character displacement through a strong reduction in average beak size. As expected, more *SS* individuals (associated with small beaks) survived, and more *LL* individuals (large beaks) died, with heterozygotes showing intermediate survival, consistent with an additive genetic effect (Fig. 2F). The frequency of the *S* allele was 61% and 37% among those that survived and those that died, respectively ($P = 0.005$, Fisher's exact test, two-sided), with a selection coefficient against *LL* homozygotes as high as $s = 0.59 \pm 0.15$. A linear regression analysis indicated that the shift in allele frequency at this locus explains about 30% of the phenotypic shift in beak size due to natural selection (*12*). Within genotypic classes, survival was nonrandom. Individuals with small beaks survived better than those with large beaks among the *LL* homozygotes ($F_{1,18} = 4.9$, $P = 0.04$) and among heterozygotes ($F_{1,30} = 10.1$, $P = 0.003$). *SS* homozygotes showed no significant association ($F_{1,17} = 0.55$, $P = 0.47$), probably because so few individuals died ($n = 5$). Thus, we conclude that the relationship between *HMGA2* and fitness was mediated entirely by the effect of this locus on beak size or associated craniofacial bones or muscles; developmental research will be necessary to reveal the underlying mechanism for the association. There is no evidence of pleiotropic effects of the gene on other, unmeasured, traits affecting fitness (table S5). Survivors were smaller in body size (*11*), but our analysis provides no additional insight into the genetic basis of body size variation (Fig. 2E) (*12*).

Introgressive hybridization can increase genetic variation and facilitate or enhance an evolutionary response to selection and adaptation (*20*, *21*), but the actual genes conferring a selective advantage are rarely known (*7*, *22*). Previous field studies have documented rare but recurring introgressive hybridization on Daphne Major between medium ground finches and small ground finch immigrants (*23*). Although the sample sizes are small, it appears that the *HMGA2*S* allele is fixed in the small ground finch ($n = 14$; fig. S2A). Positive selection for the *S* allele suggests that introgression from the small ground finch contributed to the genetic response to directional selection and character displacement in the medium ground finch.

Our results provide evidence of two loci with major effects on beak morphology across Darwin's finches. *ALX1*, a transcription factor gene, has been associated with beak shape (*15*), and here we find that *HMGA2* is associated with beak size. *ALX1* and *HMGA2* are 7.5 Mb apart on chromosome 1 in chicken and zebra finch, and probably also in Darwin's finches, as expected on the basis of the very high degree of conserved synteny among birds (*24*). Beak size and beak shape are involved in all the major evolutionary shifts in the adaptive radiation of Darwin's finches (*1*). They are also subject to strong selection in contemporary time. In the character displacement episode discussed above, beak size was subject to strong directional selection: The standardized selection differential of –0.66 for sexes combined is an exceptionally high value. We have shown that the *HMGA2* locus played a critical role in this character shift. The selection coefficient at the *HMGA2* locus ($s = 0.59 \pm 0.14$) is comparable in magnitude to the selection differential on the phenotype and is higher than other examples of strong selection, such as loci associated with coat color in mice ($s < 0.42$) (*25*). The main implication of our findings is that a single locus facilitates rapid diversification. The lack of recombination between the two *HMGA2* haplotypes, together with abundant polygenic variation and ecological opportunity (*2*, *5*), may help to explain rapid speciation in this young adaptive radiation (*1*).

REFERENCES AND NOTES

1. P. R. Grant, B. R. Grant, *How and Why Species Multiply: The Radiation of Darwin's Finches* (Princeton Univ. Press, 2008).
2. J. B. Losos, *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles* (Univ. of California Press, 2009).
3. W. L. Brown Jr., E. O. Wilson, *Syst. Zool.* **5**, 49–64 (1956).
4. P. R. Grant, *Biol. J. Linn. Soc. London* **4**, 39–68 (1972).
5. D. Schluter, *The Ecology of Adaptive Radiation* (Oxford Univ. Press, 2000).
6. D. W. Pfennig, K. S. Pfennig, *Evolution's Edge: Competition and the Origins of Diversity* (Univ. of California Press, 2012).
7. M. E. Arnegard *et al.*, *Nature* **511**, 307–311 (2014).
8. Y. E. Stuart, J. B. Losos, *Trends Ecol. Evol.* **28**, 402–408 (2013).
9. J. A. Tobias *et al.*, *Nature* **506**, 359–363 (2014).
10. P. R. Grant, B. R. Grant, *40 Years of Evolution: Darwin's Finches on Daphne Major Island* (Princeton Univ. Press, 2014).
11. P. R. Grant, B. R. Grant, *Science* **313**, 224–226 (2006).
12. See supplementary materials on *Science* Online.
13. P. R. Grant, B. R. Grant, *Evolution* **48**, 297–316 (1994).
14. A. Abzhanov, M. Protas, B. R. Grant, P. R. Grant, C. J. Tabin, *Science* **305**, 1462–1465 (2004).
15. S. Lamichhaney *et al.*, *Nature* **518**, 371–375 (2015).
16. K. Pfannkuche, H. Summer, O. Li, J. Hescheler, P. Dröge, *Stem Cell Rev. Rep.* **5**, 224–230 (2009).
17. X. Zhou, K. F. Benson, H. R. Ashar, K. Chada, *Nature* **376**, 771–774 (1995).
18. M. N. Weedon *et al.*, *Nat. Genet.* **40**, 575–583 (2008).
19. G. Fatemifar *et al.*, *Hum. Mol. Genet.* **22**, 3807–3817 (2013).
20. R. C. Lewontin, L. C. Birch, *Evolution* **20**, 315–336 (1966).
21. P. W. Hedrick, *Mol. Ecol.* **22**, 4606–4618 (2013).
22. K. J. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 196–201 (2015).
23. P. R. Grant, B. R. Grant, *Biol. J. Linn. Soc. London* **117**, 812–822 (2016).

Downloaded from http://science.sciencemag.org/ on April 22, 2016

24. G. Zhang *et al.*, *Science* **346**, 1311–1320 (2014).
25. C. R. Linnen *et al.*, *Science* **339**, 1312–1316 (2013).

ACKNOWLEDGMENTS

We thank A. Garcia-Dorado, P. Hedrick, and M. Pettersson for advice concerning statistical analysis, and J. Pettersson and U. Gustafsson for expert technical assistance. NSF (USA) funded the collection of material under permits from the Galápagos and Costa Rica National Parks Services and the Charles Darwin Research Station, and in accordance with protocols of Princeton University's Animal Welfare Committee. 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 grants 80576801 and 70374401. Computer resources were supplied by UPPMAX. The Illumina reads have been submitted to the short reads archive (www.ncbi.nlm.nih.gov/sra) with accession numbers PRJNA263122 and PRJNA301892. Raw tree files for constructing Fig. 1, C and D, have been submitted to the TreeBASE database with submission ID S18636 (http://purl.org/phylo/treebase/phylows/study/TB2:S18636). 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., F.H., J.B., and C.W. performed the bioinformatic analysis with contributions from M.S.A.; and L.A., S.L., B.R.G., and P.R.G. wrote the paper with input from the other authors. All authors approved the manuscript before submission.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6284/470/suppl/DC1
Materials and Methods
Supplementary Text
Tables S1 to S5
Figs. S1 and S2
References (*26*–*47*)

16 November 2015; accepted 14 March 2016
10.1126/science.aad8786

HUMAN GENOMICS

Health and population effects of rare gene knockouts in adult humans with related parents

Vagheesh M. Narasimhan,[1] Karen A. Hunt,[2]* Dan Mason,[3]* Christopher L. Baker,[4]* Konrad J. Karczewski,[5,6]* Michael R. Barnes,[7] Anthony H. Barnett,[8] Chris Bates,[9] Srikanth Bellary,[10] Nicholas A. Bockett,[2] Kristina Giorda,[11] Christopher J. Griffiths,[2] Harry Hemingway,[12,13] Zhilong Jia,[7] M. Ann Kelly,[14] Hajrah A. Khawaja,[7] Monkol Lek,[5,6] Shane McCarthy,[1] Rosie McEachan,[3] Anne O'Donnell-Luria,[5,6] Kenneth Paigen,[4] Constantinos A. Parisinos,[2] Eamonn Sheridan,[3] Laura Southgate,[2] Louise Tee,[14] Mark Thomas,[1] Yali Xue,[1] Michael Schnall-Levin,[11] Petko M. Petkov,[4] Chris Tyler-Smith,[1] Eamonn R. Maher,[15,16] Richard C. Trembath,[2,17] Daniel G. MacArthur,[5,6] John Wright,[3] Richard Durbin,[1]†‡ David A. van Heel[2]†‡

Examining complete gene knockouts within a viable organism can inform on gene function. We sequenced the exomes of 3222 British adults of Pakistani heritage with high parental relatedness, discovering 1111 rare-variant homozygous genotypes with predicted loss of function (knockouts) in 781 genes. We observed 13.7% fewer homozygous knockout genotypes than we expected, implying an average load of 1.6 recessive-lethal-equivalent loss-of-function (LOF) variants per adult. When genetic data were linked to the individuals' lifelong health records, we observed no significant relationship between gene knockouts and clinical consultation or prescription rate. In this data set, we identified a healthy *PRDM9*-knockout mother and performed phased genome sequencing on her, her child, and control individuals. Our results show that meiotic recombination sites are localized away from *PRDM9*-dependent hotspots. Thus, natural LOF variants inform on essential genetic loci and demonstrate *PRDM9* redundancy in humans.

Complete gene knockouts, typically caused by homozygous loss-of-function (LOF) genotypes, have helped researchers identify the function of many genes, predominantly through studies in model organisms and of severe Mendelian-inherited diseases in humans. However, information on the consequences of knocking out most human genes is still lacking. Naturally occurring complete gene knockouts offer the opportunity to study the effects of lifelong germline gene inactivation in living humans. A survey of LOF variants in adult humans revealed ~100 predicted LOF genotypes per individual, describing ~20 genes that carry homozygous predicted LOF alleles and hence are likely to be completely inactivated (*1*). Almost all of these homozygous genotypes were located at common variants with allele frequency >1%, in genes whose loss is likely to have weak or neutral effects on fitness and health (*1*). In contrast, rare predicted LOF genotypes were usually heterozygous and, thus, their overall effect on gene function is not known. A large exome sequencing aggregation study [conducted by the Exome Aggregation Consortium (ExAC)] of predominantly outbred individuals identified 1775 genes with homozygous predicted LOF genotypes in 60,706 individuals (*2*). Furthermore, 1171 genes with complete predicted LOF were identified in 104,220 Icelandic individuals (*3*), and modest enrichment for homozygous predicted LOF genotypes was shown in Finnish individuals (*4*). However, even in these large samples, homozygous predicted LOF genotypes tend to occur at variants of moderate allele frequency (~1%). Hence, these approaches will not readily assess knockouts in most genes, which are lacking such variants.

To identify knockouts created by rare homozygous predicted loss-of-function (rhLOF) variants, we sequenced the exomes of 3222 UK-dwelling adults of Pakistani heritage who were characterized as healthy, type 2 diabetic, or pregnant (*5*). These individuals have a high rate of parental relatedness (often through parents who are first cousins); thus, a substantial fraction of their autosomal genome occurs in long homozygous regions inferred to be identical by descent from a recent common ancestor (autozygous). We linked each person's genotype to their health care and epidemiological records, with the aims of (i) describing the properties and assessing the health effects of naturally occurring knockouts in an adult population; (ii) understanding the architecture of gene essentiality in the human genome, through the characterization of the population genetics of LOF variants; and (iii) conducting a detailed study of a *PRDM9* gene knockout, which plays a role in human meiotic recombination (*6*).

On average, 5.6% of the coding genome was autozygous, much higher than the percentage in outbred populations with European heritage (Fig. 1A and fig. S4). We identified, per individual, an average of 140.3 nonreference predicted

[1]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. [2]Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London E1 2AT, UK. [3]Bradford Institute for Health Research, Bradford Teaching Hospitals National Health Service (NHS) Foundation Trust, Bradford BD9 6RJ, UK. [4]Center for Genome Dynamics, The Jackson Laboratory, Bar Harbor, ME 04609, USA. [5]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114, USA. [6]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [7]William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London E1 2AT, UK. [8]Diabetes and Endocrine Centre, Heart of England NHS Foundation Trust and University of Birmingham, Birmingham B9 5SS, UK. [9]TPP, Mill House, Troy Road, Leeds LS18 5TN, UK. [10]Aston Research Centre for Healthy Ageing, Aston University, Birmingham B4 7ET, UK. [11]10X Genomics, 7068 Koll Center Parkway, Suite 415, Pleasanton, CA 94566, USA. [12]Farr Institute of Health Informatics Research, London NW1 2DA, UK. [13]Institute of Health Informatics, University College London, London NW1 2DA, UK. [14]School of Clinical and Experimental Medicine, University of Birmingham, Birmingham B15 2TT, UK. [15]Department of Medical Genetics, University of Cambridge and National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre, Box 238, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK. [16]Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK. [17]Faculty of Life Sciences and Medicine, King's College London, London SE1 1UL, UK.
*These authors contributed equally to this work. †These authors contributed equally to this work. ‡**Corresponding author. Email: rd@sanger.ac.uk (R.D.); d.vanheel@qmul.ac.uk (D.A.v.H.)**

Downloaded from http://science.sciencemag.org/ on April 22, 2016

A beak size locus in Darwin's finches facilitated character displacement during a drought
Sangeet Lamichhaney, Fan Han, Jonas Berglund, Chao Wang, Markus Sällman Almén, Matthew T. Webster, B. Rosemary Grant, Peter R. Grant and Leif Andersson (April 21, 2016)
Science Translational Medicine **352** (6284), 470-474. [doi: 10.1126/science.aad8786]

Editor's Summary

Linked loci and Galapagos finch size

Observations of parallel evolution in the finches of the Galapagos, including body and beak size, contributed to Darwin's theories. Lamichhaney *et al.* carried out whole-genome sequencing of 60 Darwin's finches. These included small, medium, and large ground finches as well as small, medium, and large tree finches. A genomic region containing the *HMGA2* gene correlated strongly with beak size across different species. This locus appears to have played a role in beak diversification throughout the radiation of Darwin's finches.

Science, this issue p. 470

This copy is for your personal, non-commercial use only.

Article Tools Visit the online version of this article to access the personalization and article tools:
http://science.sciencemag.org/content/352/6284/470

Permissions Obtain information about reproducing this article:
http://www.sciencemag.org/about/permissions.dtl

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2016 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

Downloaded from http://science.sciencemag.org/ on April 22, 2016