Sciencexpress Reports

Genomic Diversity and Evolution of the Head Crest in the Rock Pigeon

Michael D. Shapiro,[1]† Zev Kronenberg,[2] Cai Li,[3,4] Eric T. Domyan,[1] Hailin Pan,[3] Michael Campbell,[2] Hao Tan,[3] Chad D. Huff,[2,5] Haofu Hu,[3] Anna I. Vickrey,[1] Sandra C. A. Nielsen,[4] Sydney A. Stringham,[1] Hao Hu,[5] Eske Willerslev,[4] M. Thomas P. Gilbert,[4,6] Mark Yandell,[2] Guojie Zhang,[3] Jun Wang[3,7,8]†

[1]Department of Biology, University of Utah, Salt Lake City, UT 84112, USA. [2]Department of Human Genetics, University of Utah, Salt Lake City, UT 84112, USA. [3]BGI-Shenzhen, Shenzhen, 518083, China. [4]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark. [5]Department of Epidemiology, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA. [6]Ancient DNA Laboratory, Murdoch University, Perth, Western Australia 6150, Australia. [7]Department of Biology, University of Copenhagen, DK-1165 Copenhagen, Denmark. [8]Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, DK-1165 Copenhagen, Denmark.

†To whom correspondence should be addressed. E-mail: mike.shapiro@utah.edu (M.D.S.); wangj@genomics.org.cn (J.W.)

The geographic origins of breeds and the genetic basis of variation within the widely distributed and phenotypically diverse domestic rock pigeon (*Columba livia*) remain largely unknown. We generated a rock pigeon reference genome and additional genome sequences representing domestic and feral populations. We found evidence for the origins of major breed groups in the Middle East and contributions from a racing breed to North American feral populations. We identified the gene *EphB2* as a strong candidate for the derived head crest phenotype shared by numerous breeds, an important trait in mate selection in many avian species. We also found evidence that this trait evolved just once and spread throughout the species, and that the crest originates early in development by the localized molecular reversal of feather bud polarity.

Since the initial domestication of the rock pigeon in Neolithic times (*1*), breeders have selected striking differences in behavior, vocalizations, skeletal morphology, feather ornaments, colors, and color patterns to establish over 350 breeds (*2*). In many cases, the number and magnitude of differences among breeds are more characteristic of macroevolutionary changes than of changes within a single species (*2, 3*). Indeed, Charles Darwin was so fascinated by domestic pigeons that he repeatedly called attention to this dramatic example of diversity within a species to communicate his ideas about natural selection (*3, 4*).

The genetic architecture for many derived traits in pigeons is probably relatively simple (*5, 6*), probably more so than that for interspecific trait variation among many wild species, because breeders often focus on qualitative rather than quantitative variation; this increases the chance of identifying genes responsible for differences among breeds. Additionally, several morphological traits show similar patterns of variation in different breeds, making it possible to test whether the same or different genes underlie similar phenotypes. Despite these advantages, the pigeon is underused as a model for the molecular genetic basis of avian variation because of the paucity of genetic and genomic resources for this bird.

We examined genomic diversity, genetic structure, and phylogenetic relationships among domestic breeds and feral populations (free-living birds descended from escaped domestics) of the rock pigeon. The pigeon reference genome was sequenced from a male Danish tumbler with the Illumina HiSeq 2000 platform, and we also resequenced 40 additional *Columba livia* genomes to 8- to 26-fold coverage (38 individuals from 36 domestic breeds and two feral pigeons) (*7*). Genome-wide nucleotide diversity in the rock pigeon ($\pi = 3.6 \times 10^{-3}$) and the mutation rate estimate in the pigeon lineage (1.42×10^{-9} substitutions per site per year $\pm 2.60 \times 10^{-12}$ SE) are comparable to those of other avian species (*8, 9*). The observed heterozygosity indicates a large effective population size for the rock pigeon of $N_e \approx 521{,}000$; demographic inferences based on the allele frequency spectrum indicate that, aside from a very recent bottleneck, N_e has been remarkably stable over the past 1.5 million generations (*7*).

Patterns of linkage disequilibrium (LD) are indicative of haplotype sizes and genome-wide recombination rates and inform decisions about genetic mapping strategies. Using genotype data from the 40 resequenced *C. livia* genomes, we found that mean "useful LD" (*10*) (coefficient of determination, $r^2 > 0.3$) decays in 2.2 kb (fig. S10J). This suggests that we should expect little LD between typical pairs of genes in an analysis across breeds; thus, the pigeon is well suited for association-mapping strategies.

We leveraged our whole-genome data to determine breed relationships, using 1.48 million variable loci. A neighbor-joining tree rooted on *C. rupestris*, the sister species of *C. livia* (*11*), yielded several well-supported groups (Fig. 1 and fig. S16). Notably, the two feral pigeons grouped with the wattle and homer breeds (Fig. 1, pink branches), supporting the idea that escaped racing homers are probably major contributors to feral populations (*12*). As with many domesticated species, pigeon evolution is probably not exclusively linear or hierarchical (*12*). We therefore examined genetic structure among breeds by analyzing 3950 loci with ADMIXTURE (*13*) and found a best model fit at $K = 1$ (a single population, where K is the number of assumed ancestral populations). However, higher values of K can also be biologically informative (figs. S17 to S20). Our analysis includes some of the oldest lineages of domestic pigeons and breeds that were not exported from the Middle East until the late 19th or early 20th centuries (*14*), providing information about likely geographic origins of breeds and their exchange along ancient trade routes (*7*).

Derived traits in domesticated birds tend to evolve along a predictable temporal trajectory, with color variation appearing in the earliest stages of domestication, followed by plumage and structural (skeletal and soft tissue) variation, and finally behavioral differences (*2*). One of the genetically simplest derived traits of pigeons is the head crest. Head crests are common ornaments in many bird species (*2*) and are important display structures in mate selection (*15*). In pigeons, head crests consist of neck and occipital feathers with reversed growth polarity, so that the feathers grow toward the top of the head instead of down the neck. Crests can be as small and simple as a peak of feathers or as elaborate as the hood of the Jacobin, which envelops the head (Fig. 2A). Classical genetics experiments suggest that the head crest segregates as a simple Mendelian recessive trait (*6, 14*). Moreover, previous studies suggest that the same locus controls the presence of a crest in numerous breeds,

either with alternative alleles at this locus or additional modifier loci controlling the extent of crest development (*6*, *14*).

We resequenced eight individuals with head crests to directly test whether the same mutation controls crest development in different breeds. We sorted genomic variants from birds with and without head crests into separate bins and calculated allele frequency differentiation (F_{ST}) across the genome (Fig. 2B). We identified a region of high differentiation between crested and uncrested birds in the pigeon ortholog of *Ephrin receptor B2* (*EphB2*; F_{ST} = 0.94, top hit genome-wide; fig. S22A) (Fig. 2D). The role of *EphB2* in feather growth is not known, but it plays important roles in tissue patterning and morphogenesis and is a member of a receptor tyrosine kinase family that mediates development of the feather cytoskeleton (*16*, *17*). All eight crested birds were homozygous for a T nucleotide at scaffold 612, position 596613 (hereafter, the *cr* allele), whereas uncrested birds were heterozygous (n = 3) or homozygous (n = 30, including the uncrested outgroup *C. rupestris*) for the putatively ancestral C nucleotide (the + allele). These results were consistent with the known simple recessive architecture of the trait and implicated a common polymorphism associated with head crest development in multiple breeds with different genetic histories (Fig. 1). This trend extended well beyond our resequencing panel: We genotyped an additional 61 crested birds from 22 breeds and 69 uncrested birds from 57 breeds, and found a perfect association between *cr/cr* genotype and the crest phenotype (Fig. 2F). By treating the genomes of crested and uncrested birds as separate populations, we also found suggestive evidence for positive selection around the *cr* allele using cross-population extended haplotype homozygosity analysis (Fig. 2D and figs. S21 and S22B).

We then used the Variant Annotation, Analysis, and Search Tool [VAAST (*18*)] to investigate the pigeon genomes for additional coding changes associated with the head crest phenotype. This identified one gene with genome-wide significance: *EphB2*, and specifically the *cr* single-nucleotide polymorphism (SNP) ($P_{genome} = 2.0 \times 10^{-8}$) (Fig. 2, C and D). The *cr* allele has a predicted charge-changing arginine (basic) to cysteine (polar uncharged) transition in the catalytic loop of the intracellular tyrosine kinase domain of *EphB2* (Fig. 2E). This amino acid position is invariant among other vertebrates, suggesting strong purifying selection for conserved protein function. The same DLAARN to DLAACN motif change we observe in EphB2 is sufficient to abrogate kinase activity in human and mouse orthologs of the protein tyrosine kinase ZAP-70, and in both mammals and pigeons the mutant phenotypes are inherited recessively (*19*). Hence, the pigeon *cr* mutation probably abrogates kinase activity in EphB2 and disrupts downstream signal propagation, consistent with the high VAAST score for this gene. *EphB2* is therefore a convincing candidate for the *cr* locus of classical pigeon genetics (*5*–*7*, *14*).

In several wild and domesticated species, the repeated evolution of a derived trait has occurred by selection on the same gene, possibly due to the repeated selection on the same allele or haplotype (*20*–*22*). Similarly, the *cr* SNP is part of a 27.4-kb haplotype that is shared by all crested pigeons, suggesting that the mutation occurred just once and spread to multiple breeds by introgression among domestic breeds, or was selected repeatedly from a standing variant in wild rock pigeons (Fig. 2G and fig. S23; the core haplotype containing the *cr* mutation is reduced to 11 kb when uncrested heterozygotes are included). The only gene present in the shared *cr* haplotype is *EphB2* (Fig. 2D, green bar), although at this time we cannot rule out the presence of regulatory variants that might alter the expression of another gene. Crested members of the toy, fantail, Iranian, Jacobin, and owl breed groups are not more closely related to each other than to uncrested breeds (Fig. 1). Nevertheless, members of these groups had head crests hundreds of years ago (*14*), so some of these introgression events must have occurred in the distant past. Breeds with a wide variety of crest phenotypes share the same derived allele; therefore, allelic variation at the *cr* locus alone does not control all aspects of crest development (*14*). Other genetic and developmental factors beyond this locus must contribute to variation in crest morphology, akin to the presumed complex genetic architecture of species-level divergence in feather ornaments (*2*).

In crested pigeons, feather placode polarity and bud outgrowth are inverted during embryogenesis (Fig. 3). Expression of *EphB2* is not polarized in early placodes (fig. S26), so the effects of the *cr* mutation on feather polarity are probably exerted earlier in development. Why might the crest phenotype be limited to the head and neck? In Naked neck chicken mutants, regionalized production of retinoic acid allows uniform up-regulation of *Bmp7* expression to change skin phenotypes in the neck but not the body (*23*). Similarly, the head crests of several chicken breeds, in which feathers are elongated but do not have a reversed growth trajectory as in pigeons, are localized to the top of the head, probably due to ectopic expression of *Hox* positional cues (*24*). Together these examples provide evidence for regionalization of the developing head and neck skin in the chicken. We propose that analogous mechanisms might underlie skin regionalization in the pigeon and allow *cr* to change feather polarity in the occiput and neck, but not elsewhere.

Our study of domestic rock pigeons illustrates how combining comparative genomics and population-based analyses forwards our understanding of genetic relationships and the genomic basis of traits. Many of the traits that vary among pigeon breeds also vary among wild species of birds and other animals (*2*, *25*); thus, pigeons are a model for identifying the genetic basis of variation in traits of general interest. Moreover, variation in many traits in domestic pigeons, including the head crest phenotype described here, is constructive rather than regressive: Breeds derived from the ancestral rock pigeon possess traits that the ancestor does not have. Although adaptive regressive traits are important, the genetic basis of constructive traits in vertebrates remains comparatively poorly understood. The domestic pigeon is thus a promising model with which to explore the genetic architecture of derived, constructive phenotypes in a bird that is amenable to genetic, genomic, and developmental investigation.

References and Notes

1. C. A. Driscoll, D. W. Macdonald, S. J. O'Brien, From wild animals to domestic pets, an evolutionary view of domestication. *Proc. Natl. Acad. Sci. U.S.A.* **106** (suppl. 1), 9971 (2009). doi:10.1073/pnas.0901586106 Medline
2. T. D. Price, Domesticated birds as a model for the genetics of speciation by sexual selection. *Genetica* **116**, 311 (2002). doi:10.1023/A:1021248913179 Medline
3. C. Darwin, On the Origin of Species by Means of Natural Selection (John Murray, London, 1859).
4. C. R. Darwin, The Variation of Animals and Plants Under Domestication (John Murray, London, 1868), vol. 1.
5. T. H. Morgan, Notes on two crosses between different races of pigeons. *Biol. Bull.* **21**, 215 (1911). doi:10.2307/1536043
6. A. Sell, *Breeding and Inheritance in Pigeons* (Schober Verlags-GmbH, Hengersberg, Germany, 1994).

<foot>7. See the supplementary materials on *Science* Online.</foot>

8. C. N. Balakrishnan, S. V. Edwards, Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* **181**, 645 (2009). doi:10.1534/genetics.108.094250 Medline
9. H. Ellegren *et al*., The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* **491**, 756 (2012). Medline
10. J. Aerts *et al*., Extent of linkage disequilibrium in chicken. *Cytogenet. Genome Res.* **117**, 338 (2007). doi:10.1159/000103196 Medline
11. K. P. Johnson *et al*., A molecular phylogeny of the dove genera *Streptopelia* and *Columba*. *Auk* **118**, 874 (2001).
12. S. A. Stringham *et al*., Divergence, convergence, and the ancestry of feral populations in the domestic rock pigeon. *Curr. Biol.* **22**, 302 (2012).
13. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of

ancestry in unrelated individuals. *Genome Res.* **19**, 1655 (2009). doi:10.1101/gr.094052.109 Medline
14. W. M. Levi, *The Pigeon* (Levi Publishing, Sumpter, SC, ed. 2 revised, 1986).
15. T. Amundsen, Why are female birds ornamented? *Trends Ecol. Evol.* **15**, 149 (2000). doi:10.1016/S0169-5347(99)01800-5 Medline
16. I. W. McKinnell, H. Makarenkova, I. de Curtis, M. Turmaine, K. Patel, EphA4, RhoB and the molecular development of feather buds are maintained by the integrity of the actin cytoskeleton. *Dev. Biol.* **270**, 94 (2004). doi:10.1016/j.ydbio.2004.02.007 Medline
17. R. N. Kelsh, M. L. Harris, S. Colanesi, C. A. Erickson, Stripes and belly-spots—a review of pigment cell morphogenesis in vertebrates. *Semin. Cell Dev. Biol.* **20**, 90 (2009). doi:10.1016/j.semcdb.2008.10.001 Medline
18. M. Yandell *et al.*, A probabilistic disease-gene finder for personal genomes. *Genome Res.* **21**, 1529 (2011).
19. M. E. Elder *et al.*, Distinct T cell developmental consequences in humans and mice expressing identical mutations in the DLAARN motif of ZAP-70. *J. Immunol.* **166**, 656 (2001). Medline
20. P. F. Colosimo *et al.*, Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* **307**, 1928 (2005). doi:10.1126/science.1107239 Medline
21. N. B. Sutter *et al.*, A single IGF1 allele is a major determinant of small size in dogs. *Science* **316**, 112 (2007). doi:10.1126/science.1137045 Medline
22. A. S. Van Laere *et al.*, A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* **425**, 832 (2003). doi:10.1038/nature02064 Medline
23. C. Mou *et al.*, Cryptic patterning of avian skin confers a developmental facility for loss of neck feathering. *PLoS Biol.* **9**, e1001028 (2011). doi:10.1371/journal.pbio.1001028 Medline
24. Y. Wang *et al.*, The crest phenotype in chicken is associated with ectopic expression of *HOXC8* in cranial skin. *PLoS ONE* **7**, e34012 (2012). doi:10.1371/journal.pone.0034012 Medline
25. L. F. Baptista, J. E. Gomez Martinez, H. M. Horblit, Darwin's pigeons and the evolution of the columbiforms: Recapitulation of ancient genes. *Acta Zoologica Mex.* **25**, 719 (2009).
26. R. Li *et al.*, The sequence and de novo assembly of the giant panda genome. *Nature* **463**, 311 (2010). doi:10.1038/nature08696 Medline
27. R. Li *et al.*, SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966 (2009). doi:10.1093/bioinformatics/btp336 Medline
28. G. Benson, Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573 (1999). doi:10.1093/nar/27.2.573 Medline
29. R. A. Dalloul *et al.*, Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): Genome assembly and analysis. *PLoS Biol.* **8**, e1000475 (2010). doi:10.1371/journal.pbio.1000475 Medline
30. L. Hillier *et al.*; International Chicken Genome Sequencing Consortium, Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**, 695 (2004). doi:10.1038/nature03154 Medline
31. W. C. Warren *et al.*, The genome of a songbird. *Nature* **464**, 757 (2010). doi:10.1038/nature08819 Medline
32. E. Birney, M. Clamp, R. Durbin, GeneWise and Genomewise. *Genome Res.* **14**, 988 (2004). doi:10.1101/gr.1865504 Medline
33. C. Trapnell, L. Pachter, S. L. Salzberg, TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105 (2009). doi:10.1093/bioinformatics/btp120 Medline
34. C. Trapnell *et al.*, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511 (2010). doi:10.1038/nbt.1621 Medline
35. M. Stanke, S. Waack, Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, (Suppl 2), ii215 (2003). doi:10.1093/bioinformatics/btg1080 Medline
36. C. Burge, S. Karlin, Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78 (1997). doi:10.1006/jmbi.1997.0951 Medline
37. C. G. Elsik *et al.*, Creating a honey bee consensus gene set. *Genome Biol.* **8**, R13 (2007). doi:10.1186/gb-2007-8-1-r13 Medline
38. R. Apweiler *et al.*, UniProt: The Universal Protein knowledgebase. *Nucleic Acids Res.* **32** (database issue), D115 (2004). doi:10.1093/nar/gkh131 Medline
39. R. Apweiler *et al.*, The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res.* **29**, 37 (2001). doi:10.1093/nar/29.1.37 Medline
40. M. Kanehisa, S. Goto, KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27 (2000). doi:10.1093/nar/28.1.27 Medline
41. T. M. Lowe, S. R. Eddy, tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955 (1997). Medline
42. E. P. Nawrocki, D. L. Kolbe, S. R. Eddy, Infernal 1.0: Inference of RNA alignments. *Bioinformatics* **25**, 1335 (2009). doi:10.1093/bioinformatics/btp157 Medline
43. S. Griffiths-Jones, A. Bateman, M. Marshall, A. Khanna, S. R. Eddy, Rfam: An RNA family database. *Nucleic Acids Res.* **31**, 439 (2003). doi:10.1093/nar/gkg006 Medline
44. H. Li *et al.*, TreeFam: A curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res.* **34** (database issue), D572 (2006). doi:10.1093/nar/gkj118 Medline
45. W. Huang, B. T. Sherman, R. A. Lempicki, Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1 (2009). doi:10.1093/nar/gkn923 Medline
46. T. De Bie, N. Cristianini, J. P. Demuth, M. W. Hahn, CAFE: A computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269 (2006). doi:10.1093/bioinformatics/btl097 Medline
47. S. Guindon, F. Delsuc, J. F. Dufayard, O. Gascuel, Estimating maximum likelihood phylogenies with PhyML. *Methods Mol. Biol.* **537**, 113 (2009). doi:10.1007/978-1-59745-251-9_6 Medline
48. A. Morgulis *et al.*, Database indexing for production MegaBLAST searches. *Bioinformatics* **24**, 1757 (2008). doi:10.1093/bioinformatics/btn322 Medline
49. R Development Core Team, R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2008).
50. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289 (2004). doi:10.1093/bioinformatics/btg412 Medline
51. S. Purcell *et al.*, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559 (2007). doi:10.1086/519795 Medline
52. N. A. Rosenberg, DISTRUCT: A program for the graphical display of population structure. *Mol. Ecol. Notes* **4**, 137 (2004). doi:10.1046/j.1471-8286.2003.00566.x
53. A. R. Rogers, C. Huff, Linkage disequilibrium between loci with unknown phase. *Genetics* **182**, 839 (2009). doi:10.1534/genetics.108.093153 Medline
54. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403 (1990). Medline
55. D. Posada, K. A. Crandall, MODELTEST: Testing the model of DNA substitution. *Bioinformatics* **14**, 817 (1998). doi:10.1093/bioinformatics/14.9.817 Medline
56. M. A. Pacheco *et al.*, Evolution of modern birds revealed by mitogenomics: Timing the radiation and origin of major orders. *Mol. Biol. Evol.* **28**, 1927 (2011). doi:10.1093/molbev/msr014 Medline
57. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586 (2007). doi:10.1093/molbev/msm088 Medline
58. K. Nam *et al.*, Molecular evolution of genes in avian genomes. *Genome Biol.* **11**, R68 (2010). doi:10.1186/gb-2010-11-6-r68 Medline
59. R. N. Gutenkunst, R. D. Hernandez, S. H. Williamson, C. D. Bustamante, Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**, e1000695 (2009). doi:10.1371/journal.pgen.1000695 Medline
60. R. R. Hudson, Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* **18**, 337 (2002). doi:10.1093/bioinformatics/18.2.337 Medline
61. B. S. Weir, C. C. Cockerham, Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358 (1984). doi:10.2307/2408641
62. S. R. Browning, B. L. Browning, Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **81**, 1084 (2007). doi:10.1086/521987 Medline
63. M. Clement, D. Posada, K. A. Crandall, TCS: A computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657 (2000). doi:10.1046/j.1365-294x.2000.01020.x Medline
64. J. D. Thompson, D. G. Higgins, T. J. Gibson, CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic*

Acids Res. **22**, 4673 (1994). doi:10.1093/nar/22.22.4673 Medline
65. L. L. Abler *et al.*, A high throughput in situ hybridization method to characterize mRNA expression patterns in the fetal mouse lower urogenital tract. *J. Vis. Exp.* **2011**, e2912 (2011).
66. A. Fazl, The art of training pigeons in the East. *The Zoologist (London)* **12**, 167 (1888) [annotated translation from Ain-i-Akbari, 1590].
67. W. B. Tegetmeier, *Pigeons: Their Structure, Varieties, Habits, and Management* (George Routledge and Sons, London, 1868).
68. National Pigeon Association, *2010 National Pigeon Association Book of Standards* (Purebred Pigeon Publishing, Goodlettsville, TN, 2010).
69. K. Pelak *et al.*, The characterization of twenty sequenced human genomes. *PLoS Genet.* **6**, e1001111 (2010). doi:10.1371/journal.pgen.1001111 Medline
70. W. J. Kent, BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656 (2002). Medline
71. D. T. Jones, W. R. Taylor, J. M. Thornton, The rapid generation of mutation data matrices from protein sequences. *CABIOS* **8**, 275 (1992). Medline
72. K. Tamura *et al.*, MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731 (2011). doi:10.1093/molbev/msr121 Medline

Acknowledgments: We thank the University of Washington Burke Museum for the *C. rupestris* tissue sample (UWBM 59803); J. Oldham, K. Wright, the Utah Pigeon Club, the National Pigeon Association, and A. and H. O. Christiansen for domestic pigeon samples; D. Clayton for feral samples; and D. Clayton, M. Horvath, D. Kingsley, R. Nielsen, and W. Warren for discussion and comments. Supported by a Burroughs Wellcome Fund Career Award in the Biomedical Sciences, NSF CAREER DEB-1149160, University of Utah Research Foundation (M.D.S.); NIH training grant T32GM007464 (S.A.S.); NIH training grant T32HD07491 (E.T.D.); an NSF EDEN internship (A.I.V.); NIH/NHGRI R01HG004694 and NIH ARRA GO RC2HG005619 (M.Y.); and the Danish National Research Foundation (M.T.P.G. and E.W.). We acknowledge a computer time allocation from the Center for High Performance Computing at the University of Utah. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under accession no. AKCR00000000 (the first version described here is AKCR01000000; raw reads, SRA052637); RNA-seq data for annotation, accession no. GSE39333; and raw reads for resequenced genomes, accession no. SRA054391.

Supplementary Materials
www.sciencemag.org/cgi/content/full/science.1230422/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S27
Tables S1 to S28
References (*26*–*72*)

19 September 2012; accepted 7 December 2012
Published online 31 January 2013; 10.1126/science.1230422

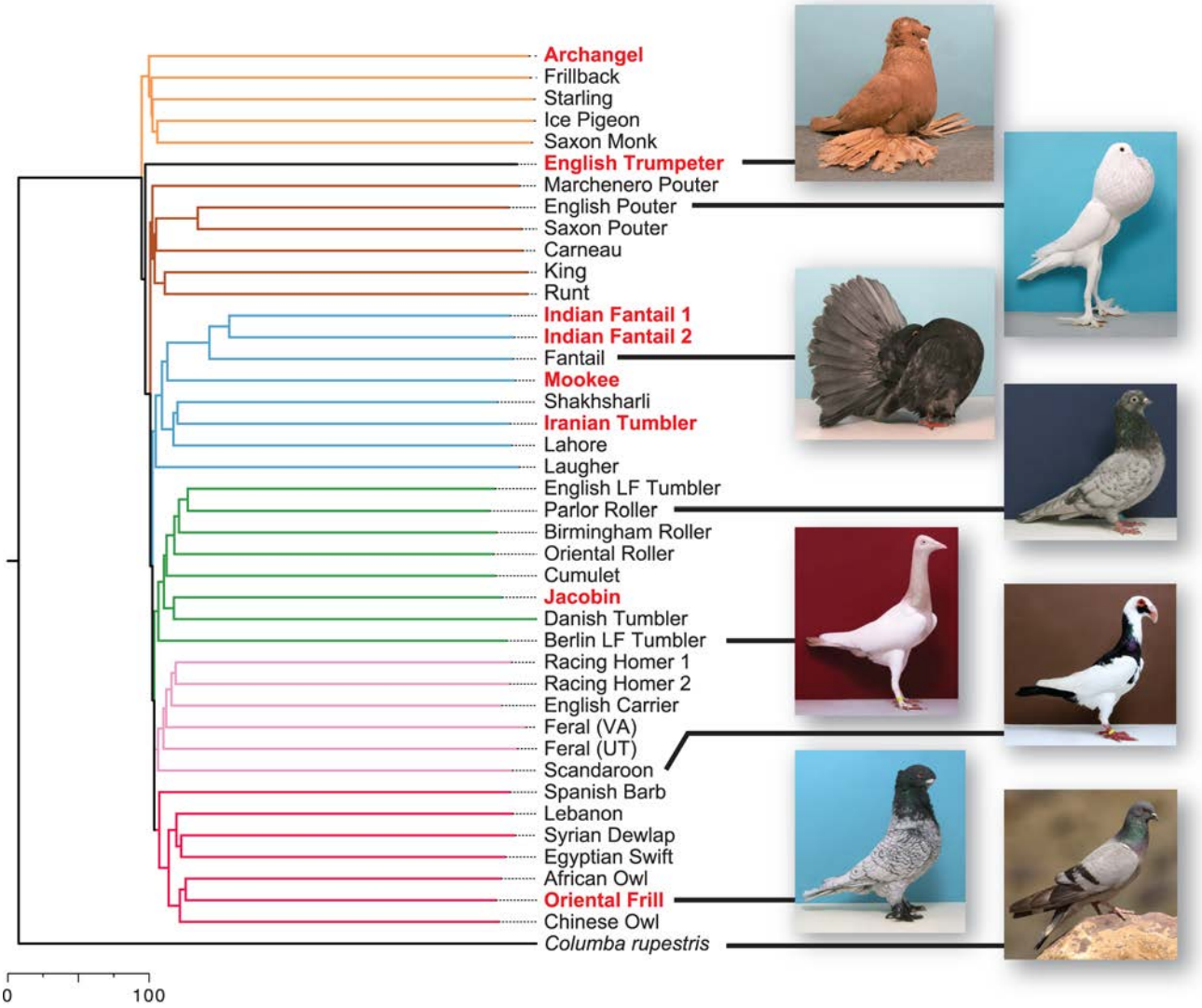

Fig. 1. Relationships among rock pigeons and the hill pigeon *C. rupestris*. A consensus neighbor-joining tree based on 1.48 million genomic SNPs and 1000 bootstrap replicates (see fig. S16 for bootstrap support) is shown. Branches are colored according to traditional breed groups (*12*) and/or geographic affinities: orange, toy breeds; brown, pouters and utility breeds; light blue, Indian and Iranian breeds; green, tumblers and highflyers; pink, homers and wattle breeds; red, Mediterranean and owl breeds; black, voice characteristics (*14*). Bold red lettering indicates breeds with the head crest phenotype. Scale bar, Euclidean distance. [Photo credits: T. Hellmann (domestic breeds) and M. V. Shreeram (*C. rupestris*)]

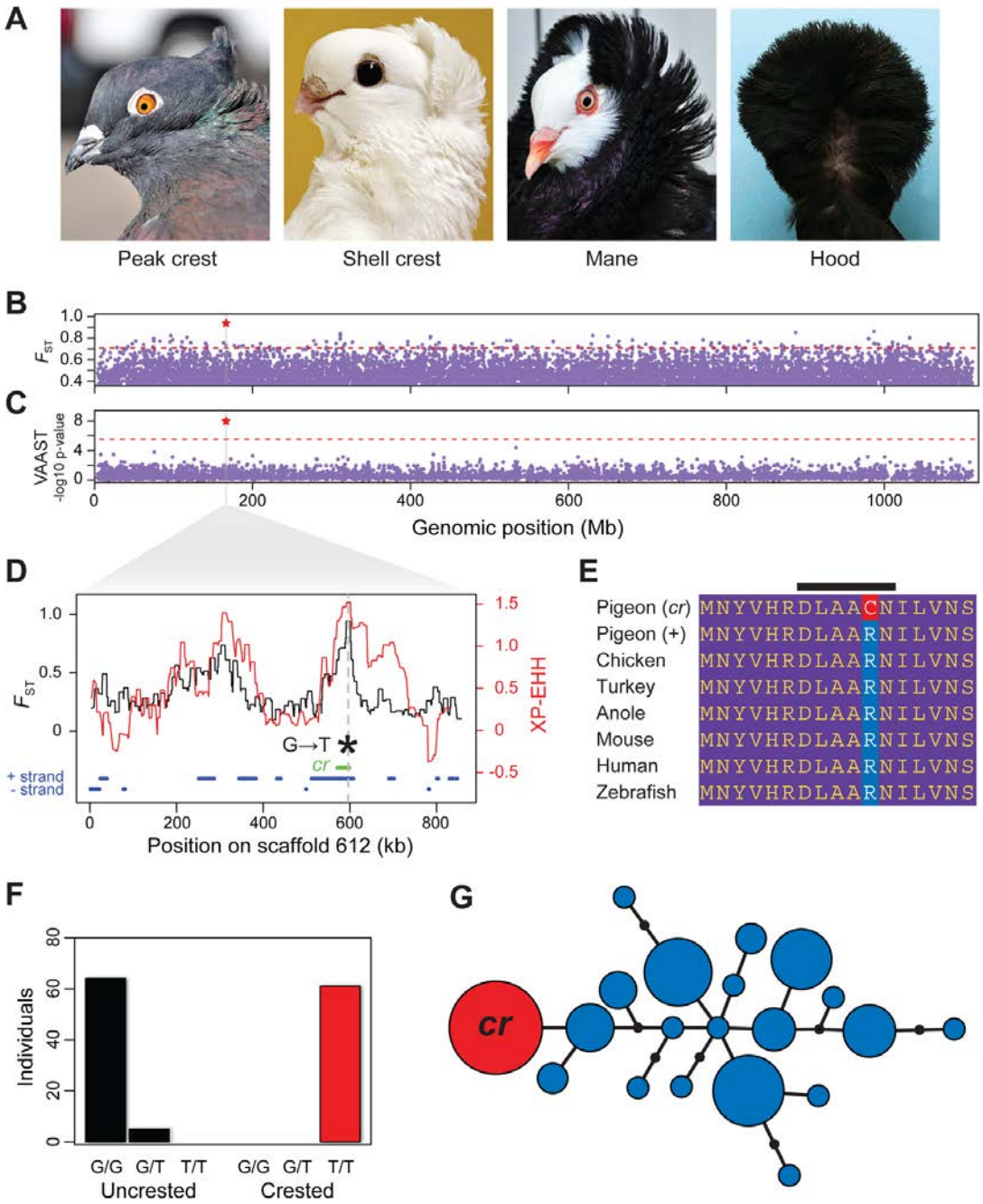

Fig. 2. *EphB2* is associated with the derived head crest phenotype. (**A**) Head crests are variable among breeds (left to right: Indian fantail, Old German owl, Old Dutch capuchin, Jacobin). (**B**) F_{ST} between crested and uncrested pigeons, with maximum value for individual SNPs plotted for nonoverlapping 100-kb windows across the genome. Red star, window with the highest score. Dashed red line, top 1% of scores. (**C**) Genome-wide VAAST scan. Each dot represents a single gene. Red star, gene with the highest score. Dashed red line, genome-wide significance cutoff. (**D**) Magnification of scaffold 612 in shaded region of (B) and (C). Black trace, maximum F_{ST} between crested and uncrested birds over a 300-SNP window. Red trace, unstandardized cross-population extended haplotype homozygosity (XP-EHH); higher values are evidence of selection (see fig. S21, genome-wide plot). Dashed vertical line, position of the lone genome-wide significant VAAST hit. Green bar, the 27.4-kb haplotype shared by all crested birds, includes only the *EphB2* gene. Blue bars, gene predictions on **+** and **–** DNA strands. (**E**) The *cr* mutation induces a charge-changing amino acid substitution; black bar, highly conserved DLAARN motif of catalytic loop. (**F**) Genotypes of 159 birds from 79 breeds at the *cr* locus are perfectly associated with the crest phenotype under a recessive model. (**G**) Network diagram of the minimal 11-kb haplotype shared by all resequenced rock pigeons with the *cr* mutation (also see fig. S23). Many haplotypes contain the + allele (blue), but only one contains the *cr* SNP (red). The sizes of the circles are proportional to the number of chromosomes containing a haplotype. Line segments represent single-nucleotide differences. [Jacobin photo credit: T. Hellmann]

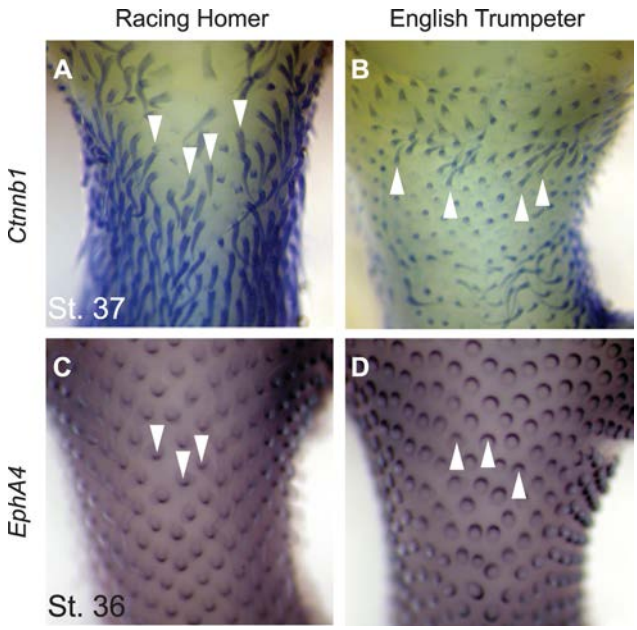

Fig. 3. Feather bud polarity is reversed in the *cr* mutant. (**A** and **B**) Expression of the feather structural gene *Ctnnb1* reveals the direction of outgrowth of early feather buds. St., Hamburger-Hamilton embryonic stage. (A) Neck and occipital head expression of *Ctnnb1* in an embryo of the uncrested racing homer. Feather buds point downward along the contour of the head and neck (arrowheads). (B) Occipital feather buds point upward in the equivalent region of the crested English trumpeter, indicating morphological reversal of feather orientation. (**C** and **D**) Expression of the polarity marker *EphA4* was assayed at an earlier developmental stage to test whether feather placodes, the ectodermal thickenings that give rise to feather buds, are also reversed. (C) Polarity marker *EphA4* is expressed posteriorly (arrowheads) in feather placodes of the racing homer. (D) The polarity of placodes is reversed in the English trumpeter. Expression of *EphB2* in the skin is weak and unpolarized at this stage in both morphs (fig. S26).