The genomic landscape underlying phenotypic integrity in the face of gene flow in crows

J. W. Poelstra *et al.*

Science **344**, 1410 (2014);
DOI: 10.1126/science.1253226

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.

Diseases (NIAID) grant R01AI091823 to R.J.; by the NIH grant 8P51OD011103-51 (to the New England Primate Research Center), 1R56AI097834-01, and 1R01NS080833-01 to M.D.; and by the Swiss Federal Office for Civil Protection BABS #353003325 to A.R. NE-CAT at the Advanced Photon Source (APS) is supported by a grant from the National Institute of General Medical Sciences (P41 GM103403). Use of the APS, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under contract DE-AC02-06CH11357. The atomic coordinates and structure factors of the mini-HA–EC1–EC2 complex have been deposited in the Protein Data Bank under the accession code 4QD2. BoNT availability is subject to the restrictions that apply to Centers for Disease Control and Prevention select agents and NIAID Category A pathogens. R.J. is named as an inventor on a U.S. patent application applied for by University of California, Irvine related to oral drug delivery systems.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/344/6190/1405/suppl/DC1
Materials and Methods
Figs. S1 to S9
Table S1
References (*29*–*42*)

24 March 2014; accepted 27 May 2014
10.1126/science.1253823

ECOLOGICAL GENOMICS

The genomic landscape underlying phenotypic integrity in the face of gene flow in crows

J. W. Poelstra,[1]* N. Vijay,[1]* C. M. Bossu,[1]* H. Lantz,[2,3] B. Ryll,[4] I. Müller,[5,6] V. Baglione,[7] P. Unneberg,[8] M. Wikelski,[5,6] M. G. Grabherr,[3] J. B. W. Wolf[1]†

The importance, extent, and mode of interspecific gene flow for the evolution of species has long been debated. Characterization of genomic differentiation in a classic example of hybridization between all-black carrion crows and gray-coated hooded crows identified genome-wide introgression extending far beyond the morphological hybrid zone. Gene expression divergence was concentrated in pigmentation genes expressed in gray versus black feather follicles. Only a small number of narrow genomic islands exhibited resistance to gene flow. One prominent genomic region (<2 megabases) harbored 81 of all 82 fixed differences (of 8.4 million single-nucleotide polymorphisms in total) linking genes involved in pigmentation and in visual perception—a genomic signal reflecting color-mediated prezygotic isolation. Thus, localized genomic selection can cause marked heterogeneity in introgression landscapes while maintaining phenotypic divergence.

Genomic studies increasingly appreciate the importance of interspecific gene flow in the context of species diversification (*1*, *2*), including that of hominids (*3*). Yet our understanding of the forces generating heterogeneity in differentiation across the genome and their relationship to phenotypic evolution is limited (*3*–*5*). The European hybrid zone between all-black carrion crows [*Corvus* (*corone*) *corone*] and gray-coated hooded crows [*C.* (*corone*) *cornix*] (*6*) provides ideal conditions for studying the evolutionary consequences of introgression in a phase of early species divergence. The geographical distribution pattern of these species suggests a population history shaped by glaciation cycles during the Pleistocene, when periods of isolation in distinct southern refugia alternated with periods of range expansion and secondary contact (*7*). The near-absence of neutral genetic differentiation across the hybrid zone (*8*) strongly contrasts with abrupt spatial segregation of plumage coloration (Fig. 1), which has remained remarkably stable over the past century despite successful backcrossing of hybrids (*6*). Evidence for color-assortative mating (*9*) and the independent recurrence of pied plumage phenotypes in the otherwise all-black genus *Corvus* has prompted the hypothesis that this color polymorphism is promoting speciation (fig. S1).

We assembled a high-quality reference genome of one hooded crow male sequenced to 152× coverage (table S1). The final assembly consisted of 1298 scaffolds with an N50 size of 16.4 Mb (*10*) and a cumulative length of 1.04 Gb, including 3.52% gaps; 95% of the assembly was contained within the 100 largest scaffolds (size range: 1 to 50 Mb), which approached chromosome size (crow karyotype: $2n$ = 80; size range of zebra finch chromosomes: 1 to 156 Mb) (fig. S2, A and B, and table S2). Evidence-based annotation of the genome with mRNA sequencing data (table S3) identified a set of 20,794 protein-coding genes.

We subsequently resequenced 60 genomes of unrelated individuals from two populations each of carrion and hooded crows (Fig. 1) to a mean sequence coverage of 12.2× (range: 7.1× to 28.6×, table S4); 8.44 million single-nucleotide polymorphisms (SNPs) segregated across all four populations, 5.27 million of which were shared between carrion and hooded crows. Several lines of evidence suggest substantial genome-wide gene flow across the hybrid zone. The major axes of genetic variation (accounting for 11.5% of the total variance) coincided with the presumed direction of spatial expansion out of Spain, yet German carrion crows clustered more closely with both hooded crow populations [Fig. 1; for pairwise fixation index (F_{ST}) estimates see Fig. 2A, fig. S3, and table S5; see fig. S4 for mitochondrial DNA divergence]. Extensive gene flow between hooded crows and the German carrion crow population was further supported by ABBA-BABA tests (fig. S5 and table S6), admixture analyses (fig. S6), and coalescence-based parameter estimates of an isolation-with-migration model (fig. S7). Consistent with a scenario of admixture upon secondary contact, all populations but Spain showed signatures of expansion and had higher nucleotide diversity than that of the presumably refugial Spanish population (table S5).

The low genome-wide differentiation across the hybrid zone was also reflected in gene expression divergence as measured by mRNA sequencing (RNA-seq) from 19 individuals and five tissues under controlled conditions. The overall proportion of differentially expressed genes between carrion and hooded crows was low, between 0.03% and 0.41% across tissues [Fig. 3B; see (*8*); false discovery rate <0.05]. The main difference in gene expression patterns corresponded to phenotypic divergence in plumage coloration. Differentially expressed genes predominated in growing feather follicles from the torso, where carrion crows produce black feathers and hooded crows gray feathers (Figs. 1 and 3A; $\chi^2_{df=4}$ = 169.34, P < 0.001), with an overrepresentation of genes in the melanogenesis pigmentation pathway (Fig. 3, B and C; n = 21, $\chi^2_{df=4}$ = 147.16, P < 0.001; see table S7 for Gene Ontology terms). Nineteen of 20 differentially expressed melanogenesis-related genes were underexpressed in the gray hooded crow torso (table S8). The pathway-wide pattern of reduced expression could not be attributed to differences in melanocyte density (Fig. 4), but rather to upstream regulatory changes inducing broad-scale down-regulation in follicles producing gray feathers (Fig. 3C).

We surveyed the landscape of genomic divergence with a window-based approach based on population genetic summary statistics derived for 22,072 nonoverlapping 50-kb windows (Fig. 2A and fig. S8) and a hypothesis-free clustering

[1]Science for Life Laboratory and Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, 75236 Uppsala, Sweden. [2]Bioinformatics Infrastructure for Life Sciences, Uppsala University, 75124 Uppsala, Sweden. [3]Science for Life Laboratory and Department of Medical Biochemistry and Microbiology, Uppsala University, 75123 Uppsala, Sweden. [4]Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, 75236 Uppsala, Sweden. [5]Max Planck Institute for Ornithology, 78315 Radolfzell, Germany. [6]Department of Biology, University of Konstanz, 78464 Konstanz, Germany. [7]Departamento de Ciencias Agro-Forestales, Campus La Yutera, Universidad de Valladolid, 34004 Palencia, Spain. [8]Science for Life Laboratory and Department of Biochemistry and Biophysics, Stockholm University, 171 21 Solna, Sweden. *These authors contributed equally to this work. †Corresponding author. E-mail: jochen.wolf@ebc.uu.se

algorithm reconstructing local genomic phylogenies (so-called "cacti") (*10*) (Fig. 2C). Phylogenetic clustering segmented the genome into two major topologies (class I and class II cacti hereafter; fig. S9), with only 0.28% of the genome following a topology that clearly separated carrion and hooded crows. Both methods identified one 1.95-Mb genomic region exhibiting extreme genetic differentiation between carrion and hooded crows on scaffolds 78 and 60, which, on the basis of synteny with other birds, are located adjacently on chromosome 18 [Fig. 2, A and B, and fig. S8A; for a detailed exploration of divergence peaks, see (*10*) and fig. S8]. This region, harboring 81 of all 82 fixed sites between carrion and hooded crows, contained 40 annotated protein-coding genes, had strongly reduced levels of nucleotide diversity, and showed increased linkage disequilibrium (LD). Disrupted patterns of LD, taxon-specific dropout of paired-end read mapping (*10*) (fig. S10), and the presence of two local F_{ST} peaks connected by a saddle (*11*) suggest the possibility of an inversion in this region.

In the center of the right local F_{ST} peak (Fig. 2B), a region about 4.1 kb upstream of the *PRKCA* gene exhibited longer haplotype blocks (iES statistic) and showed evidence for recent, hooded crow–specific, positive selection: It was enriched for fixed hooded crow–specific derived variants (22 of 28) and had significantly reduced values of Fu and Li's D statistic ($P < 0.05$) in hooded crows. Furthermore, this region contains a tandem array of voltage-dependent calcium channel subunits, including those encoded by *CACNG5*, *CACNG4*, and *CACNG1*. The *CACNG* gene family codes for transmembrane regulators of AMPA receptors, which influence the microphthalmia-associated transcription factor gene *MITF*, a central regulatory element of the melanogenesis pathway (*12*) (Fig. 3C). At least 11 of the 21 melanogenesis genes underexpressed in gray hooded crow feather follicles are directly regulated by *MITF* (e.g., *TYR*, *TYRP1*, *SLC45A2*, *RAB38*, *EDNRB2*) or interact with *MITF* in feedback loops (*MC1R*, *HPGDS*) (Fig. 3C and table S8). Moreover, *RAS-GRF1*, a central mediator between calcium concentration and downstream gene expression (*13*), was significantly underexpressed in the gray feather follicles of hooded crows (Fig. 3C and table S8). This tentatively links gene expression, color phenotype, and the signature of divergent selection in the divergence peak.

Two additional strongly differentially expressed melanogenesis genes were located in divergent genomic regions: *NDP* (Norrie disease pseudoglioma) and *HPGDS* (hematopoietic prostaglandin D synthase), which also act closely upstream of *MITF* (*14*) (Fig. 3C and fig. S11). *HPGDS* was among the most strongly underexpressed genes in the hooded crow torso and also showed an extreme reduction in variance, consistent with selection (Ansari-Bradley test, $P = 0.0039$). A multigenic architecture of the color trait is consistent with the segregation pattern of different crow color phenotypes in the hybrid zone (fig. S12), indicating that a number of genes likely contribute to color divergence.

Under the left local F_{ST} peak, fixed differences enriched for carrion crow–derived variants were clustered in a narrowly defined area (Fig. 2B). An exception included a hooded-derived fixed variant in the regulatory region of the gene *RGS9* (regulator of G protein signaling 9), which affects visual perception in vertebrates (*15*), was highly expressed in eye (higher than in any other tissue by a factor of 5), and showed lower expression in hooded crows across tissues (fig. S11). Not only is *RGS9* involved in phototransduction, but different splice forms (also detected in crows; fig. S11) regulate dopamine and opioid signaling in the brain (*15*). This potential pleiotropy is of interest because the main differences between carrion and hooded crows are coloration, color-assortative mate choice (*9*), and hormone-mediated social dominance behavior (*16*). These diverse biological functions may also be explained by the *SCL24* gene family (*17*): *SLC24A2* and *SLC24A3*, involved in vision and hormonal regulation, respectively, were associated with class I cacti in crows (Fig. 2C), and *SLC24A4*, involved in pigmentation,

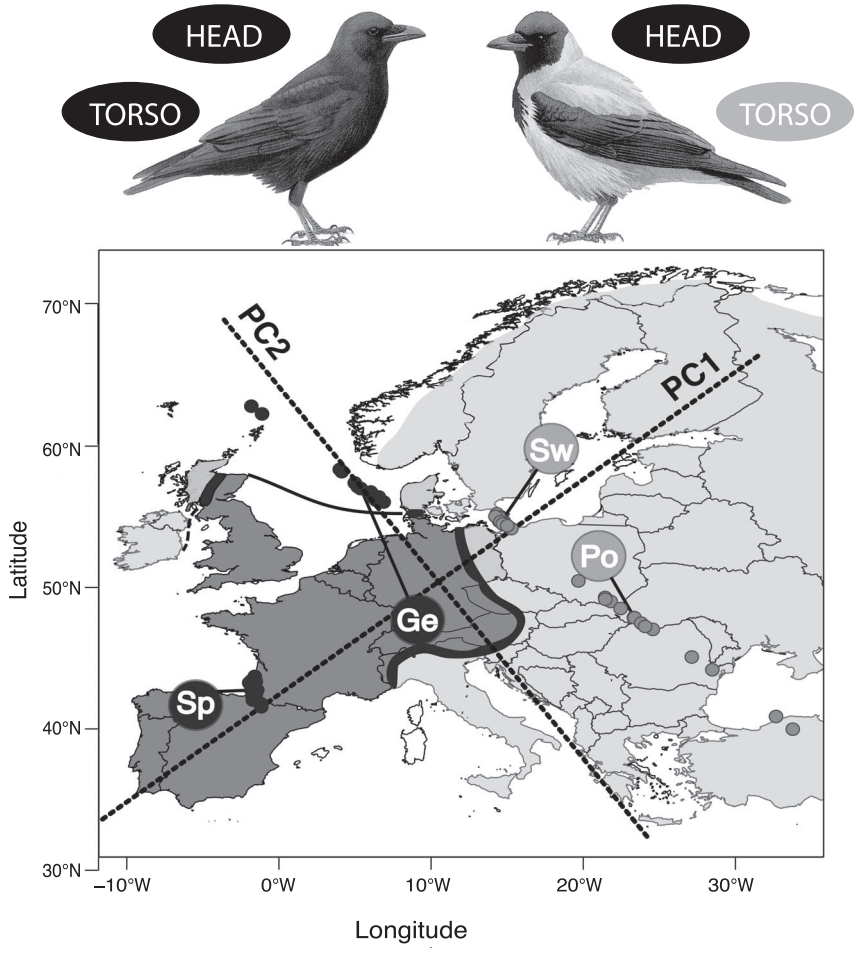

Fig. 1. The crow system. Illustrations of carrion crow (left) and hooded crow (right). We sampled gene expression data of regrowing feathers from the head and torso, which differ in coloration. The map below shows the European distribution of the carrion crow (dark gray area) and hooded crow (light gray area); the hybrid zone is shown as a black solid line. Sampling locations of this study in Spain (Sp), Germany (Ge), Poland (Po), and Sweden (Sw) are indicated by filled circles (black, carrion crow; gray, hooded crow). Superimposed on the geographical map are the major axes of genetic variation (principal components PC1 and PC2) derived from whole-genome resequencing data from the four populations (black and gray dots, respectively). [Drawings courtesy of Dan Zetterström]

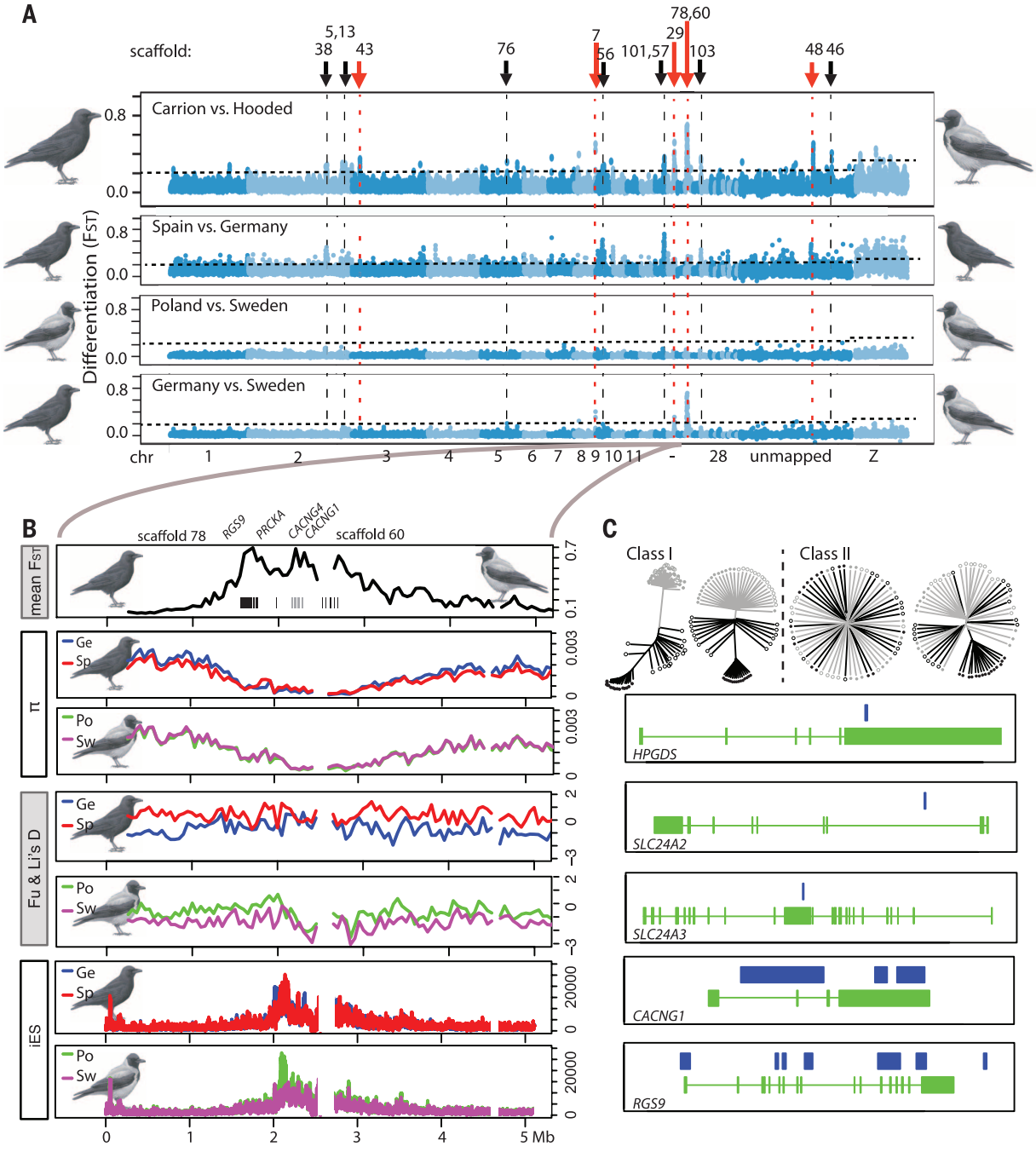

Fig. 2. The genomic landscape of divergence. (**A**) Pairwise genetic differentiation (F_{ST}) in 50-kb sliding windows across the genome between carrion crows and hooded crows (top panel), between carrion crow populations (second panel), between hooded crow populations (third panel), and between carrion crows and hooded crows from across the hybrid zone (bottom panel). Alternating colors denote the different chromosomes; and the dotted horizontal lines mark the 99th-percentile F_{ST} estimates of autosomes and the Z chromosome, respectively. Red arrows indicate taxon-specific peaks of consecutive elevated F_{ST} windows; black arrows highlight outlier peaks that are specific to comparisons with the refugial Spanish population. (**B**) The largest and most extreme F_{ST} peak was located on scaffolds 78 and 60. Plotted within the peak are 81 of 82 total fixed variants (black bars in top panel, ancestral alleles; gray bars, derived hooded crow–specific alleles). (**C**) Localized phylogenetic patterns within the genome ("cacti"). The majority of the genome exhibits neutral genetic divergence classified as class II cacti (divergent Spanish population or complete admixture). Two cacti display clear separation of carrion crows (black: Spain, solid circles; Germany, open circles) and hooded crows (gray: Sweden, solid circles, Poland, open circles) (class I cacti; fig. S9). Four exemplar gene models associated with the melanogenesis pathway and associated with visual perception are depicted (green blocks and lines represent exons and introns, respectively) along with the regions assigned to class I cacti (blue).

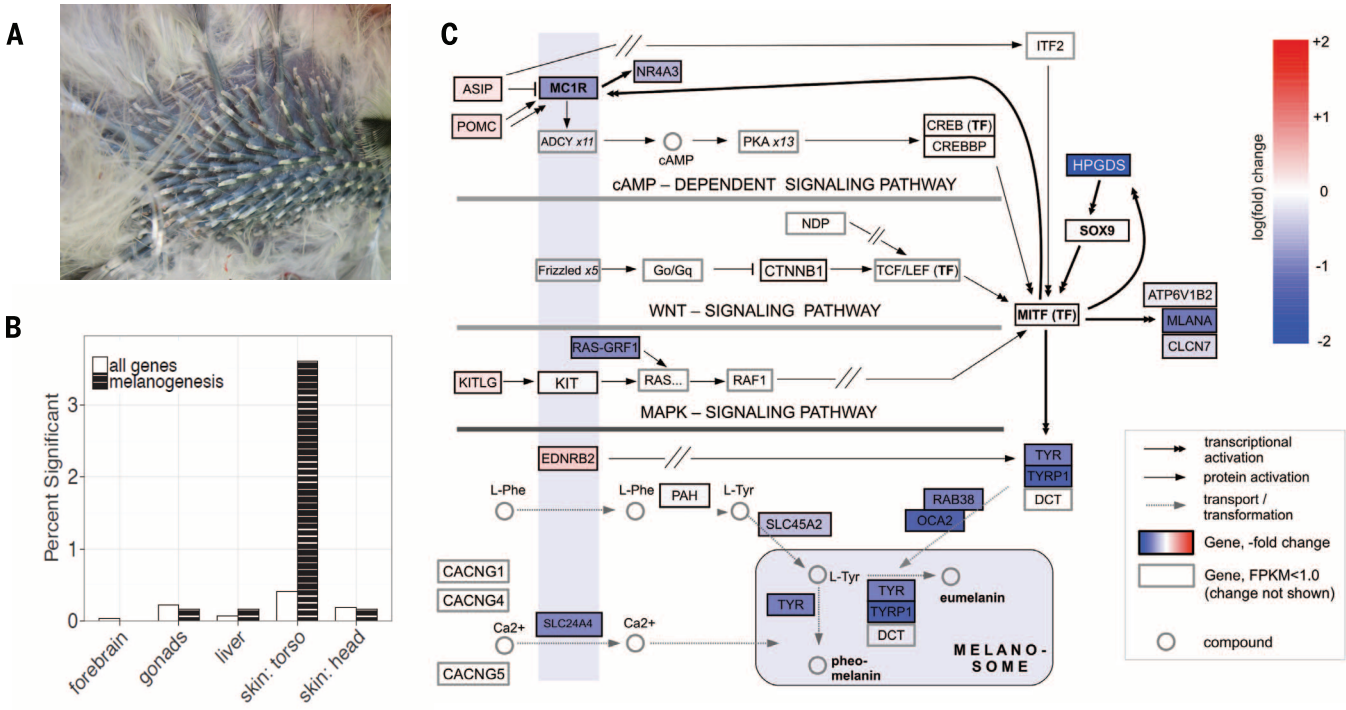

Fig. 3. The functional genomic basis of plumage color differences. (**A**) Photograph of the developmental stage of feather follicles used for gene expression quantification. (**B**) Percentage of all expressed genes (white) and melanogenesis genes (striped) inferred to be differentially expressed. (**C**) Schematic overview of the melanogenesis pathway, with genes underexpressed in skin from torso in hooded crows shown in blue, genes overexpressed shown in red (color scale indicates relative change). Genes either not supported by RNA-seq data or expressed at FPKM (fragments per kilobase of exon per million fragments mapped) levels below 1.0 in both hooded and carrion crows are shown as gray boxes. Pathway relationships are shown for transcriptional activation (black double arrows), protein activation (e.g., by phosphorylation) (black single arrows), and transport (gray single arrow, dotted line).

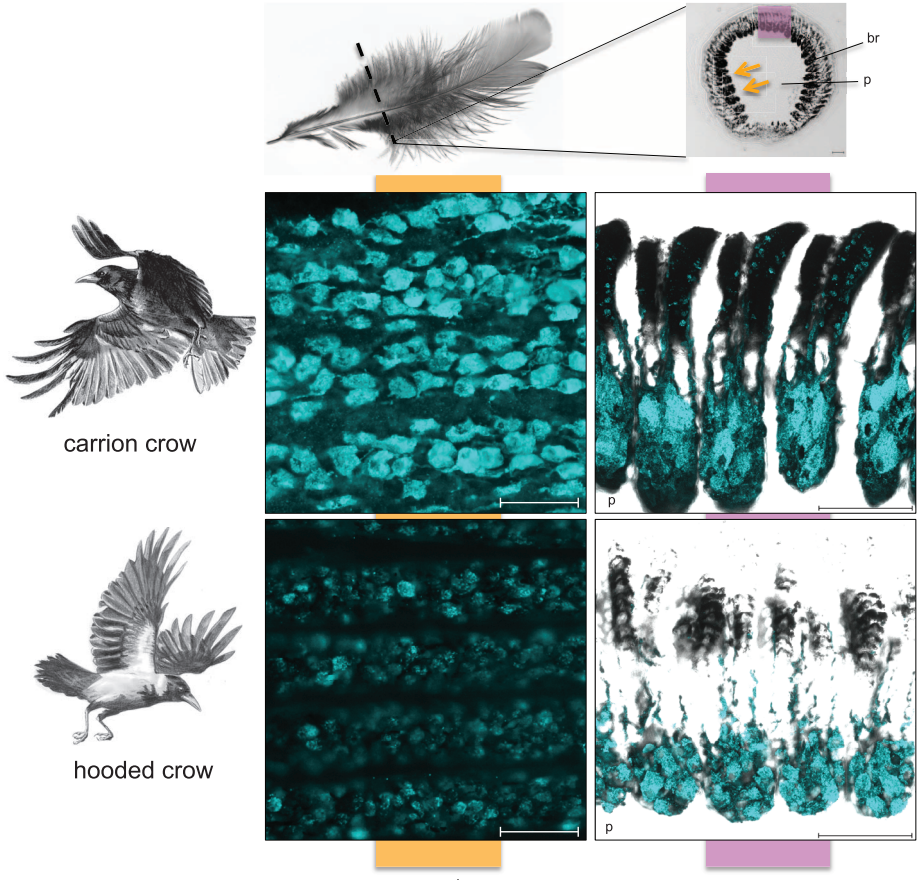

Fig. 4. Characterization of feather melanocytes. Immunohistochemistry on feather follicles of carrion crows (upper panel) and hooded crows (lower panel) for one of the melanogenic enzymes (a-Trp1, cyan) differentially expressed in our RNA-seq data confirms differential melanogenesis pathway activity (left column; view on to barbed ridges of whole-mount feather bulges, yellow arrows) but shows no striking differences in melanocyte density (right column; detail of sectioned feather follicle, magenta square). Matched samples from carrion crows and hooded crows were stained in parallel and imaged under identical conditions to ensure comparability. br, barbed ridge; p, pulp. Scale bar, 50 μm. [Drawing courtesy of Kristina Fraune; feather photo from Wikimedia Commons]

was strongly underexpressed in hooded crows (Fig. 3C).

Our results show that the pronounced phenotypic differences between hooded and carrion crows are maintained in the face of substantial gene flow by variation in less than 1% of the genome. The crow system thus constitutes an intriguing test case for the appropriate level of taxonomic delineation (*18*) and the interpretation of local peaks of divergence (often called "speciation islands") (*19*). Gene flow slightly reduced the width of the main divergence peak and eroded linkage disequilibrium at the peak margins (fig. S10B) but slightly increased the peak amplitude (*10*). Hence, divergent selection acting on small genomic regions did not appear to expand local differentiation or seed genome-wide reproductive isolation (*19*), although it did maintain phenotypic differentiation in the face of gene flow. Our findings further highlight the role of inversions in evolution (*20*, *21*), potentially linking key genes involved in a mating trait and a tentative visual preference locus. This establishes an interesting parallel to studies in which single genes (*22*), "supergenes" (*4*), or large-scale inversions (*23*) have been implicated in driving morphological change relevant to mate choice, but contrasts with studies in which many ecologically relevant loci (*21*) or structural genomic features dominated the genomic landscape of divergence (*24*). A key feature that distinguishes the crow system is the apparent lack of ecological selection on the maintenance of separate phenotypes. Instead, the data presented here are consistent with the idea that assortative mating and sexual selection can exclusively cause phenotypic and genotypic differentiation (*25*).

REFERENCES AND NOTES

1. R. Abbott *et al*., *J. Evol. Biol.* **26**, 229–246 (2013).
2. S. H. Martin *et al*., Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* **23**, 1817–1828 (2013).
3. S. Sankararaman *et al*., *Nature* **507**, 354–357 (2014).
4. Heliconius Genome Consortium, *Nature* **487**, 94–98 (2012).
5. H. Ellegren *et al*., *Nature* **491**, 756–760 (2012).
6. W. Meise, *J. Ornithol.* **76**, 1–203 (1928).
7. G. Hewitt, *Nature* **405**, 907–913 (2000).
8. J. B. W. Wolf *et al*., *Mol. Ecol.* **19** (suppl. 1), 162–175 (2010).
9. C. Randler, *Ardea* **95**, 143–149 (2007).
10. See supplementary materials on *Science* Online.
11. R. F. Guerrero, F. Rousset, M. Kirkpatrick, *Philos. Trans. R. Soc. London Ser. B* **367**, 430–438 (2012).
12. M. J. Hoogduijn *et al*., *Pigment Cell Res.* **19**, 58–67 (2006).
13. C. L. Farnsworth *et al*., *Nature* **376**, 524–527 (1995).
14. B. Moniot *et al*., *Dev. Dyn.* **240**, 2335–2343 (2011).
15. K. A. Martemyanov, V. Y. Arshavsky, *Prog. Mol. Biol. Transl. Sci.* **86**, 205–227 (2009).
16. N. Saino, L. Scatizzi, *Boll. Zool.* **58**, 255–260 (1991).
17. P. P. M. Schnetkamp, *Mol. Aspects Med.* **34**, 455–464 (2013).
18. D. T. Parkin, M. Collinson, A. J. Helbig, A. G. Knox, G. Sangster, *Br. Birds* **96**, 274–290 (2003).
19. J. L. Feder, S. P. Egan, P. Nosil, *Trends Genet.* **28**, 342–350 (2012).
20. A. A. Hoffmann, L. H. Rieseberg, *Annu. Rev. Ecol. Evol. Syst.* **39**, 21–42 (2008).
21. F. C. Jones *et al*., *Nature* **484**, 55–61 (2012).
22. K. Kunte *et al*., *Nature* **507**, 229–232 (2014).
23. H. Thorneycroft, *Evolution* **29**, 611–621 (1975).
24. S. Renaut *et al*., *Nat. Commun.* **4**, 1827 (2013).
25. C. A. McLean, D. Stuart-Fox, *Biol. Rev. Camb. Philos. Soc.* 10.1111/brv.12083 (2014).

ACKNOWLEDGMENTS

We thank W. Fiedler, K. Temnow, L. Vossen, M. Döpfner, T. Pärt, K.-H. Siebenrock, and the other staff of the Max Planck Institute for Ornithology for help with field work and animal husbandry. R. Burri assisted in population genetic analyses. Supported by VolkswagenStiftung grant I/83 496, Swedish Research Council grant 621-2010-5553, the Wallenberg Advanced Bioinformatics Infrastructure (funded by the Knut and Alice Wallenberg Foundation), Carl Tryggers foundation grant CTS11:517, CTS12:543, and European Research Council grant ERCStG-336536 (J.B.W.W.); the Max Planck Society (M.W., I.M.); and SciLifeLab Uppsala (M.G.G., H.L.). *C. brachyrhynchos* samples were provided by courtesy of the American Museum of Natural History, New York; Polish hooded crows by A. Kruszewicz and M. Gajownik, Warsaw Zoo. We thank the UPPMAX Next-Generation Sequencing Cluster and Storage (UPPNEX) project, funded by the Knut and Alice Wallenberg Foundation and the Swedish National Infrastructure for Computing, for access to computational resources. All sequencing data and the genome assembly are available in the NCBI short read archive under project PRJNA192205.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/344/6190/1410/suppl/DC1
Supplementary Text
Figs. S1 to S13
Tables S1 to S10
References (*26*–*97*)

11 March 2014; accepted 5 May 2014
10.1126/science.1253226