Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication

Miguel Carneiro *et al.*

Science **345**, 1074 (2014);
DOI: 10.1126/science.1253714

closure (*17*); and channel opening occurs when LBD clamshells adopt their maximally closed conformation (*22*, *23*, *30*–*32*), represented by agonist-bound structures of isolated LBD (*13*, *17*) with the D1-D1 interface intact (fig. S7). The first model represents a traditional view (*18*) where the final desensitized state has the D1-D1 interface modified (fig. S12A). In this model, the $GluA2_{NOW}$ structure represents the agonist-bound closed state, which is predicted to be a transient state with negligible occupancy (Fig. 1B and fig. S2F) insufficient to produce protein crystals. Nevertheless, such a scenario is plausible if only a limited range of conformations of the protein is accessible in the solubilized receptor or the crystal lattice contacts substantially affect protein conformation.

The second model (fig. S12B) assumes two-step desensitization with $GluA2_{NOW}$ representing a deep desensitized state. This model is consistent with the predictions of kinetic modeling that, at high NOW concentrations, the majority of receptors accumulate in the deep desensitized state (D_24 in fig. S2). It also predicts that the same tension forces, applied from ATD and the ion channel through the connecting linkers that open LBD clamshells during deactivation, help transition the receptor from the deep desensitized state back to the desensitized state. Therefore, the second model explains why mutations that change the rate of deactivation often produce similar effects on the rate of recovery from desensitization (*14*, *33*, *34*).

Independent of gating model, the entry into desensitization is associated with modification of the D1-D1 interface (fig. S12) (*18*, *20*, *35*–*37*). One possible modification is represented by structures of the S729C and G725C cross-linked isolated LBDs (*18*) where the D1-D1 interface is ruptured. However, K493C cross-linking does not affect desensitization (Fig. 4C) and argues against these structures representing the desensitized state of the intact receptor. Alternatively, the D1-D1 interface modification might be a rotation of the D1 lobes relative to each other that does not change the distance between K493 lysines but introduces relative displacement of pairs of other residues at the D1-D1 interface. Correspondingly, mutations like L483Y (*28*) or S497C (Fig. 4C and fig. S11) or positive allosteric modulators like CTZ (*10*, *29*) would block desensitization by imposing constraints on the D1-D1 interface rearrangement.

REFERENCES AND NOTES

1. S. F. Traynelis *et al.*, *Pharmacol. Rev.* **62**, 405–496 (2010).
2. D. Bowie, *CNS Neurol. Disord. Drug Targets* **7**, 129–143 (2008).
3. P. Paoletti, C. Bellone, Q. Zhou, *Nat. Rev. Neurosci.* **14**, 383–400 (2013).
4. Z. Galen Wo, R. E. Oswald, *Trends Neurosci.* **18**, 161–168 (1995).
5. A. I. Sobolevsky, M. P. Rosconi, E. Gouaux, *Nature* **462**, 745–756 (2009).
6. C. H. Lee *et al.*, *Nature* **511**, 191–197 (2014).
7. E. Karakas, H. Furukawa, *Science* **344**, 992–997 (2014).
8. M. V. Jones, G. L. Westbrook, *Trends Neurosci.* **19**, 96–101 (1996).
9. M. D. Black, *Psychopharmacology (Berl.)* **179**, 154–163 (2005).
10. D. K. Patneau, L. Vyklicky Jr., M. L. Mayer, *J. Neurosci.* **13**, 3496–3509 (1993).
11. R. Jin, T. G. Banke, M. L. Mayer, S. F. Traynelis, E. Gouaux, *Nat. Neurosci.* **6**, 803–810 (2003).
12. K. Poon, L. M. Nowak, R. E. Oswald, *Biophys. J.* **99**, 1437–1446 (2010).
13. K. Poon, A. H. Ahmed, L. M. Nowak, R. E. Oswald, *Mol. Pharmacol.* **80**, 49–59 (2011).
14. A. Robert, N. Armstrong, J. E. Gouaux, J. R. Howe, *J. Neurosci.* **25**, 3752–3762 (2005).
15. R. Jin *et al.*, *EMBO J.* **28**, 1812–1823 (2009).
16. K. Menuz, R. M. Stroud, R. A. Nicoll, F. A. Hays, *Science* **318**, 815–817 (2007).
17. N. Armstrong, E. Gouaux, *Neuron* **28**, 165–181 (2000).
18. N. Armstrong, J. Jasti, M. Beich-Frandsen, E. Gouaux, *Cell* **127**, 85–97 (2006).
19. A. S. Maltsev, A. H. Ahmed, M. K. Fenwick, D. E. Jane, R. E. Oswald, *Biochemistry* **47**, 10600–10610 (2008).
20. A. J. Plested, M. L. Mayer, *J. Neurosci.* **29**, 11912–11923 (2009).
21. M. K. Fenwick, R. E. Oswald, *J. Biol. Chem.* **285**, 12334–12343 (2010).
22. C. F. Landes, A. Rambhadran, J. N. Taylor, F. Salatan, V. Jayaraman, *Nat. Chem. Biol.* **7**, 168–173 (2011).
23. A. Y. Lau, B. Roux, *Nat. Struct. Mol. Biol.* **18**, 283–287 (2011).
24. P. A. Postila, M. Ylilauri, O. T. Pentikäinen, *J. Chem. Inf. Model.* **51**, 1037–1047 (2011).
25. S. Ramaswamy *et al.*, *J. Biol. Chem.* **287**, 43557–43564 (2012).
26. A. H. Ahmed *et al.*, *J. Biol. Chem.* **288**, 27658–27666 (2013).
27. Y. Yao, J. Belcher, A. J. Berger, M. L. Mayer, A. Y. Lau, *Structure* **21**, 1788–1799 (2013).
28. Y. Stern-Bach, S. Russo, M. Neuman, C. Rosenmund, *Neuron* **21**, 907–918 (1998).
29. Y. Sun *et al.*, *Nature* **417**, 245–253 (2002).
30. W. Zhang, Y. Cho, E. Lolis, J. R. Howe, *J. Neurosci.* **28**, 932–943 (2008).
31. A. H. Ahmed, S. Wang, H. H. Chuang, R. E. Oswald, *J. Biol. Chem.* **286**, 35257–35266 (2011).
32. D. M. MacLean, A. Y. Wong, A. M. Fay, D. Bowie, *J. Neurosci.* **31**, 2136–2144 (2011).
33. M. C. Weston, C. Gertler, M. L. Mayer, C. Rosenmund, *J. Neurosci.* **26**, 7650–7658 (2006).
34. A. L. Carbone, A. J. Plested, *Neuron* **74**, 845–857 (2012).
35. C. R. Midgett, A. Gill, D. R. Madden, *Front. Mol. Neurosci.* **4**, 56 (2012).
36. A. Y. Lau *et al.*, *Neuron* **79**, 492–503 (2013).
37. D. M. Schauder *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 5921–5926 (2013).

ACKNOWLEDGMENTS

We thank the personnel at beamlines X4A, X4C, X25, and X29 of the National Synchrotron Light Source and at beamlines 24-ID-C and 24-ID-E of the Advanced Photon Source. 24-ID-C and 24-ID-E are the Northeastern Collaborative Access Team beamlines, which are supported by a grant from the National Institute of General Medical Sciences (P41 GM103403) from the NIH. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. We thank L. Wollmuth and R. Kazi for help in setting up electrophysiological experiments and K. Saotome for comments on the manuscript. This work was supported by the NIH (NS083660) and the Klingenstein Foundation (A.I.S.). Coordinates and structure factors have been deposited in the Protein Data Bank with accession numbers 4U4F for $GluA2_{NOW}$ and 4U4G for $GluA2_{ZK}$.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6200/1070/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S3
References (*38*–*50*)

27 May 2014; accepted 25 July 2014
Published online 7 August 2014;
10.1126/science.1256508

EVOLUTIONARY GENOMICS

Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication

Miguel Carneiro,[1]* Carl-Johan Rubin,[2]* Federica Di Palma,[3,4]* Frank W. Albert,[5]† Jessica Alföldi,[3] Alvaro Martinez Barrio,[2] Gerli Pielberg,[2] Nima Rafati,[2] Shumaila Sayyab,[6] Jason Turner-Maier,[3] Shady Younis,[2,7] Sandra Afonso,[1] Bronwen Aken,[8,9] Joel M. Alves,[1,10] Daniel Barrell,[8,9] Gerard Bolet,[11] Samuel Boucher,[12] Hernán A. Burbano,[5]‡ Rita Campos,[1] Jean L. Chang,[3] Veronique Duranthon,[13] Luca Fontanesi,[14] Hervé Garreau,[11] David Heiman,[3] Jeremy Johnson,[3] Rose G. Mage,[15] Ze Peng,[16] Guillaume Queney,[17] Claire Rogel-Gaillard,[18] Magali Ruffier,[8,9] Steve Searle,[8] Rafael Villafuerte,[19] Anqi Xiong,[20] Sarah Young,[3] Karin Forsberg-Nilsson,[20] Jeffrey M. Good,[5,21] Eric S. Lander,[3] Nuno Ferrand,[1,22]* Kerstin Lindblad-Toh,[2,3]*§ Leif Andersson[2,6,23]*§

The genetic changes underlying the initial steps of animal domestication are still poorly understood. We generated a high-quality reference genome for the rabbit and compared it to resequencing data from populations of wild and domestic rabbits. We identified more than 100 selective sweeps specific to domestic rabbits but only a relatively small number of fixed (or nearly fixed) single-nucleotide polymorphisms (SNPs) for derived alleles. SNPs with marked allele frequency differences between wild and domestic rabbits were enriched for conserved noncoding sites. Enrichment analyses suggest that genes affecting brain and neuronal development have often been targeted during domestication. We propose that because of a truly complex genetic background, tame behavior in rabbits and other domestic animals evolved by shifts in allele frequencies at many loci, rather than by critical changes at only a few domestication loci.

Domestication of animals (that is, the evolution of wild species into tame forms) has resulted in notable changes in behavior, morphology, physiology, and reproduction (*1*). The genetic underpinnings of the initial steps of animal domestication are poorly understood but probably involved changes in behavior that allowed the animals to survive and reproduce under conditions that might be too stressful for wild animals. Given the differences

in behavior between wild and domesticated animals, we investigated to what extent this process involved fixation of new mutations with large phenotypic effects, as opposed to selection on standing variation. Such studies are hampered in most domestic animals due to ancient domestication events, extinct wild ancestors, or geographically widespread wild ancestors.

Rabbit domestication was initiated in monasteries in southern France as recently as ~1400 years ago (*2*). At this time, wild rabbits were mostly restricted to the Iberian Peninsula, where two subspecies occurred (*Oryctolagus cuniculus cuniculus* and *O. c. algirus*), and to France, colonized by *O. c. cuniculus* (Fig. 1B). Additionally, the area of origin of domestic rabbits is still populated with wild rabbits related to the ancestors of the domestic rabbit (*3*). This recent and well-defined origin provides a major advantage for inferring genetic changes underlying domestication.

We performed Sanger sequencing and assembly of a female rabbit genome (*4*). The draft OryCun2.0 assembly size is 2.66 Gb, with a contig N50 size of 64.7 kb and a scaffold N50 size of 35.9 Mb (tables S1 and S2). The genome assembly was annotated using the Ensembl gene annotation pipeline (Ensembl release 73, September 2013) and with both rabbit RNA sequencing data and the annotation of human orthologs (*4*) (table S3). Our analysis of rabbit domestication used Ensembl annotations as well as a custom pipeline for annotation of untranslated regions (UTRs) (168,286 distinct features), noncoding RNA (n = 9666), and noncoding conserved elements (2,518,476 distinct features).

To identify genomic regions under selection during domestication, we performed whole-genome resequencing (10× coverage) of pooled samples (table S4) of six different breeds of domestic rabbits (Fig. 1A), 3 pools of wild rabbits from southern France, and 11 pools of wild rabbits from the Iberian Peninsula, representing both subspecies (Fig. 1B). We also sequenced a close relative, the snowshoe hare (*Lepus americanus*), to deduce the ancestral state at polymorphic sites. Short sequence reads were aligned to our assembly; single-nucleotide polymorphism (SNP) calling resulted in the identification of 50 million high-quality SNPs and 5.6 million insertion/deletion polymorphisms after filtering (table S5). The numbers of SNPs at noncoding conserved sites and in coding sequences were 719,911 and 154,489, respectively. The per-site nucleotide diversity (π) within populations of wild rabbits was in the range of 0.6 to 0.9% (Fig. 1C). Thus, the rabbit is one of the most polymorphic mammals sequenced so far, presumably due to a larger long-term effective population size relative to other sequenced mammals (*5*). Identity scores confirm that the domestic rabbit is most closely related to wild rabbits from southern France (fig. S1A), and we inferred a strong correlation (r = 0.94) in allele frequencies at most loci between these groups (fig. S1B). The average nucleotide diversity of each sequenced population is consistent with a bottleneck and reduction in genetic diversity when rabbits from the Iberian Peninsula colonized southern France and a second bottleneck during domestication (*3*) (Fig. 1, B and C).

Selective sweeps occur when beneficial genetic variants increase in frequency due to positive selection together with linked neutral sequence variants (*6*). This results in genomic islands of reduced heterozygosity and increased differentiation between populations around the selected site. We compared genetic diversity between domestic rabbits as one group to wild rabbits representing 14 different locations in France and the Iberian Peninsula. We calculated fixation index (F_{ST}) values between wild and domestic rabbits and average pooled heterozygosity (H) in domestic rabbits in 50-kb sliding windows across the genome (hereafter referred to as the F_{ST}-H outlier approach). We identified 78 outliers with $F_{ST} > 0.35$ and $H < 0.05$ (Fig. 2A, fig. S2, database S1). We also used SweepFinder (*7*), which calculates maximum composite likelihoods for the presence of a selective sweep, taking into account the background pattern of genetic variation in the data and with a significance threshold set by coalescent simulations incorporating the recent demographic history of domestic and wild rabbits (figs. S3 and S4 and databases S1 and S2) (*4*). This analysis resulted in the identification of 78 significant sweeps (false discovery rate = 5%) (Fig. 2A, database S1). Thirty-one (40%) of these were also detected with the F_{ST}-H approach (Fig. 2A). This incomplete overlap is probably explained by the fact that SweepFinder primarily assesses the distribution of genetic diversity within the selected population, whereas the F_{ST}-H analysis identifies the most differentiated regions of the genome between wild and domestic rabbits. We carried out an additional screen using targeted sequence capture on an independent sample of individual French wild and domestic rabbits. We targeted more than 6 Mb of DNA sequence split into 5000 1.2-kb intronic fragments that were distributed across the genome and selected independently of the genome resequencing results above. Coalescent simulations, using the targeted resequencing data set and incorporating the recent demographic history of domestic rabbit as a null model (figs. S3 and S4 and databases S1 and S2) (*4*), revealed that the majority of the sweep regions detected by whole-genome resequencing showed levels and patterns of genetic variation that were observed less than 5% of the time in the simulated data set (76.0% with SweepFinder and 73.7% with F_{ST}-H outlier regions, excluding regions without targeted fragments), a highly significant overlap (Fisher's exact test, $P < 1 \times 10^{-5}$ for both tests). Furthermore, 26 of the 31 sweep regions detected with both SweepFinder and the F_{ST}-H approach were targeted in the capture experiment, and an even greater proportion (88.5%) showed levels and patterns of genetic variation unlikely to be generated under the specified demographic model.

An example of a selective sweep overlapping the 3′-part of *GRIK2* (glutamate receptor, ionotropic, kainate 2) is shown in Fig. 2B. Parts of this region have low heterozygosity in domestic rabbits, and at position chr12:90,153,383 base pairs, domestic rabbits carry a nearly fixed derived allele at a site with 100% sequence conservation among 29 mammals except for the allele present in domestic rabbits (*8*), suggesting functional importance. *GRIK2* encodes a subunit of a glutamate receptor that is highly expressed in the brain and has been associated with recessive mental retardation in humans (*9*). Both SweepFinder and the F_{ST}-H outlier analysis identified two sweeps near *SOX2* (SRY-BOX 2), separated by a region of high heterozygosity (Fig. 2C). *SOX2* encodes a transcription factor that is required for stem cell maintenance (*10*).

Given the comprehensive sampling in our study and the correlation in allele frequencies between domestic and French wild rabbits (fig. S1B), highly differentiated individual SNPs are likely either to have been directly targeted by selection or to occur in the vicinity of loci under selection. For each SNP, we calculated the absolute allele frequency difference between wild and domestic rabbits (ΔAF) and sorted these into 5% bins (ΔAF = 0 to 0.05, etc.). The majority of SNPs showed low ΔAF between wild and domestic

[1]CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Universidade do Porto, 4485-661, Vairão, Portugal. [2]Science for Life Laboratory Uppsala, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden. [3]Broad Institute of Harvard and Massachusetts Institute of Technology, 7 Cambridge Center, Cambridge, MA 02142, USA. [4]Vertebrate and Health Genomics, The Genome Analysis Centre, Norwich, UK. [5]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. [6]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden. [7]Department of Animal Production, Ain Shams University, Shoubra El-Kheima, Cairo, Egypt. [8]Wellcome Trust Sanger Institute, Hinxton, UK. [9]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK. [10]Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK. [11]Institut National de la Recherche Agronomique (INRA), UMR1388 Génétique, Physiologie et Systèmes d'Elevage, F-31326 Castanet-Tolosan, France. [12]Labovet Conseil, BP539, 85505 Les Herbiers Cedex, France. [13]INRA, UMR1198 Biologie du Développement et Reproduction, F-78350 Jouy-en-Josas, France. [14]Department of Agricultural and Food Sciences, Division of Animal Sciences, University of Bologna, 40127 Bologna, Italy. [15]Laboratory of Immunology, National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health, Bethesda, MD 20892, USA. [16]U.S. Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA. [17]ANTAGENE, Animal Genomics Laboratory, Lyon, France. [18]INRA, UMR1313 Génétique Animale et Biologie Intégrative, F- 78350, Jouy-en-Josas, France. [19]Instituto de Estudios Sociales Avanzados, (IESA-CSIC) Campo Santo de los Mártires 7, Córdoba, Spain. [20]Science for Life Laboratory, Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [21]Division of Biological Sciences, The University of Montana, Missoula, MT 59812, USA. [22]Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s/n. 4169-007 Porto, Portugal. [23]Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843-4458, USA.

*These authors contributed equally to this work. †Present address: Department of Human Genetics, University of California, Los Angeles, Gonda Center, 695 Charles E. Young Drive South, Los Angeles, CA 90095, USA. ‡Present address: Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany. §Corresponding author. E-mail: kersli@broadinstitute.org (K.L.-T.); leif.andersson@imbim.uu.se (L.A.)

rabbits (Fig. 2D). We examined exons, introns, UTRs, and evolutionarily conserved sites for enrichment of SNPs with high ΔAF, as would be expected under directional selection on many independent mutations (Fig. 2D and table S6). We observed no consistent enrichment for high ΔAF SNPs in introns, but we found significant enrichments in exons, UTRs, and conserved noncoding sites (χ^2 test, $P < 0.05$). We detected a significant excess of SNPs at conserved noncoding sites for each bin $\Delta AF > 0.45$ (χ^2 test, $P = 1.8 \times 10^{-3}$ to 7.3×10^{-17}), whereas in coding sequence, a significant excess was found only at $\Delta AF > 0.80$ (χ^2 test, $P = 3.0 \times 10^{-2}$ to 1.0×10^{-3}). Compared to the relative proportions in the entire data set, there was an excess of 3000 SNPs at conserved noncoding sites with $\Delta AF > 0.45$, whereas for exonic SNPs with $\Delta AF > 0.80$, the excess was only 83 SNPs (table S6). Thus, changes at regulatory sites have played a much more prominent role in rabbit domestication, at least numerically, than changes in coding sequences.

We selected the 1635 SNPs at conserved noncoding sites with $\Delta AF > 0.80$, which represent 681 nonoverlapping 1-Mb blocks of the rabbit genome. So as not to inflate significances due to inclusion of SNPs in strong linkage disequilibrium, we selected only one SNP per 50 kb, leaving 1071 SNPs. More than 60% of the SNPs were located 50 kb or more from the closest transcriptional start site (TSS) (Fig. 2E), suggesting that many differentiated SNPs are located in long-range regulatory elements. A gene ontology (GO) overrepresentation analysis (*11*) examining all genes located within 1 Mb from high-ΔAF SNPs showed that the most enriched categories of biological processes involved "cell fate commitment" (Bonferroni $P = 3.1 \times 10^{-3}$ to 5.4×10^{-5}) (Table 1 and database S3), whereas the statistically most significant categories involved brain and nervous system cell development (Bonferroni $P = 2.9 \times 10^{-3}$ to 3.7×10^{-10}). Many of the mouse orthologs of genes associated with noncoding high-ΔAF SNPs were expressed in the brain or sensory organs, and this enrichment was highly significant (Table 1). We also examined phenotypes observed in mouse mutants (www.informatics.jax.org) for these genes, revealing a significant (Bonferroni $P = 3.7 \times 10^{-2}$ to 7.5×10^{-17}) enrichment of genes associated with defects in brain and neuronal development, development of sensory organs (hearing and vision), ectoderm development, and respiratory system phenotypes (fig. S5). These highly significant overrepresentations were obtained because there were many genes in the overrepresented categories (Table 1). For example, we observed high-ΔAF SNPs associated with 191 genes (113 expected by chance) from the nervous system–development GO category (Bonferroni $P = 3.7 \times 10^{-10}$). Thus, rabbit domestication must have a highly polygenic basis with many loci responding to selection and where genes affecting brain and neuronal development have been particularly targeted.

None of the coding SNPs that differed between wild and domestic rabbits was a nonsense or frame-shift mutation, consistent with data from chicken (*12*) and pigs (*13*), suggesting that gene loss has not played a major role during animal domestication. This is an important finding, as it has been suggested that gene inactivation could be an important mechanism for rapid evolutionary change, for instance, during domestication (*14*). Of 69,985 autosomal missense mutations, there were no fixed differences, and only 14 showed a ΔAF above 90%. On the basis of poor sequence conservation, similar chemical properties of the substituted amino acids, and/or the derived state of the domestic allele, we assume that most of these result from hitchhiking rather than being functionally important (database S4). However, two missense mutations stand out; these may be direct targets of selection because at these two positions the domestic rabbit differs from all other sequenced mammals (>40 species). The first is a $Gln^{813} \rightarrow Arg^{813}$ substitution in *TTC21B* (tetratricopeptide repeat domain 21B protein), which is part of the ciliome and modulates sonic hedgehog signaling during embryonic development (*15*). The other is an $Arg^{1627} \rightarrow Trp^{1627}$ substitution in *KDM6B* (lysine-specific demethylase 6B), which encodes a histone H3K27 demethylase involved in *HOX* gene regulation during development (*16*).

Deletions distinct to domestic rabbits were difficult to identify because the genome assembly is based on a domestic rabbit, but some convincing duplications were detected with marked frequency differences between wild and domestic rabbits (database S5). We observed a

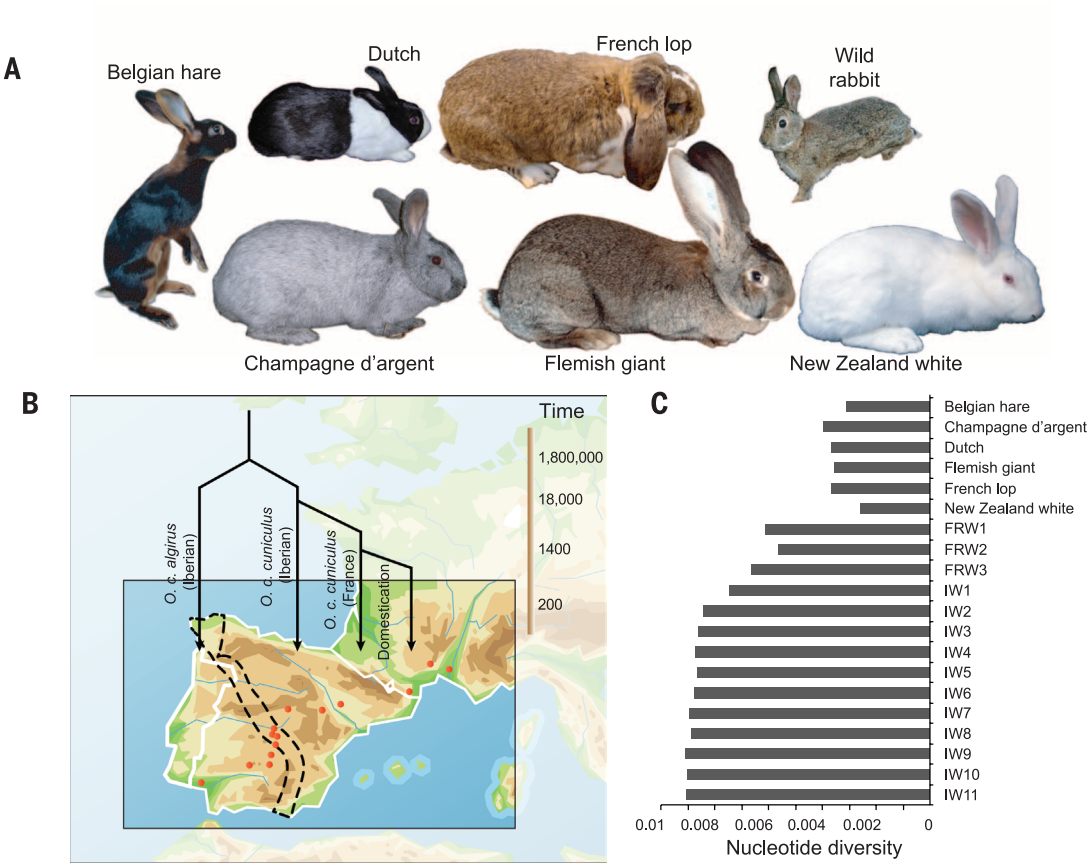

Fig. 1. Experimental design and population data. (**A**) Images of the six rabbit breeds included in the study (sized to reflect differences in body weight) and of a wild rabbit. (**B**) Map of the Iberian Peninsula and southern France with sample locations marked (orange dots). Demographic history of this species is indicated, and a logarithmic time scale is shown at right. The hybrid zone between the two subspecies is marked with dashes. (**C**) Nucleotide diversities in domestic and wild populations. The French (FRW1 to FRW3) and Iberian (IW1 to IW11) wild rabbit populations are ordered according to a northeast-to-southwest transection. Sample locations are provided in table S4.

one–base pair insertion/deletion polymorphism located within an intron of *IMMP2L* (inner mitochondrial membrane peptidase-2 like protein), where domestic and wild rabbits were fixed for different alleles. The polymorphism occurs in a sweep region and is the sequence polymorphism with highest ΔAF in the region (fig. S6). Mutations in *IMMP2L* have been associated with Tourette syndrome and autism in humans (*17*).

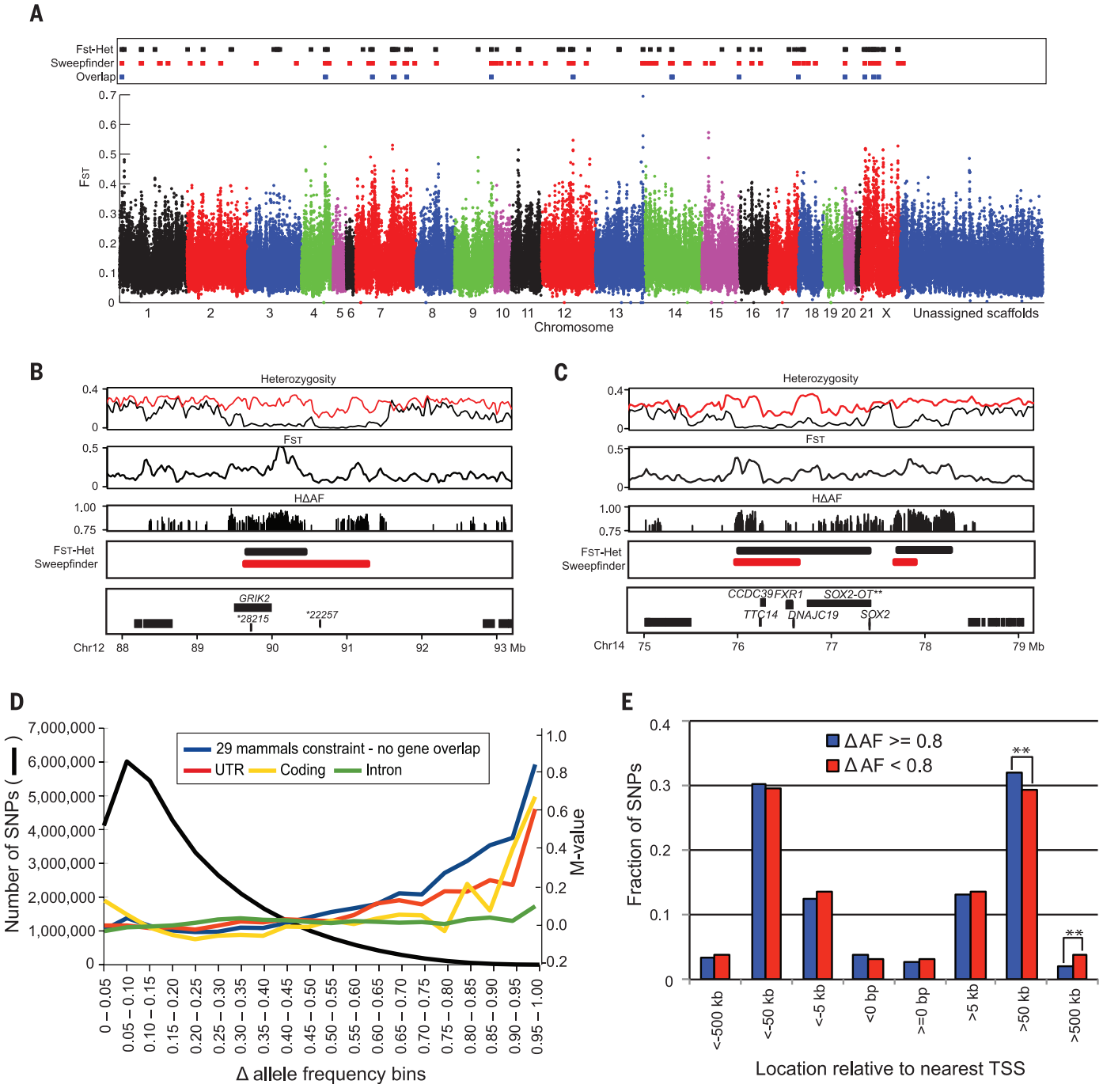

Fig. 2. Selective sweep and Δ allele frequency analyses. (**A**) Plot of F_{ST} values between wild and domestic rabbits. Sweeps detected with the F_{ST}-*H* outlier approach, SweepFinder, and their overlaps are marked on top. Unassigned scaffolds were not included in the analysis. (**B** and **C**) Selective sweeps at *GRIK2* (B) and *SOX2* (C). Heterozygosity plots for wild (red) and domestic (black) rabbits together with plots of F_{ST} values and SNPs with $\Delta AF > 0.75$ (HΔAF). The bottom panels show putative sweep regions, detected with the F_{ST}-*H* outlier approach and SweepFinder, marked with horizontal bars. Gene annotations in sweep regions are indicated: * represents ENSOCUT000000; ***SOX2-OT* represents the manually annotated SOX2 overlapping transcript (*4*). (**D**) The majority of SNPs showed low ΔAF between wild and domestic rabbits. The black line indicates the number of SNPs in nonoverlapping ΔAF bins (left *y* axis). Colored lines denote M values (log2-fold changes) of the relative frequencies of SNPs at noncoding evolutionary conserved sites (blue), in UTRs (red), exons (yellow), and introns (green), according to ΔAF bins (right *y* axis). M values were calculated by comparing the frequency of SNPs in a given annotation category in a specific bin with the corresponding frequency across all bins. (**E**) Location of SNPs at conserved noncoding sites with $\Delta AF \geq 0.8$ SNPs ($n = 1635$) and $\Delta AF < 0.8$ SNPs ($n = 502{,}343$) in relation to the TSS of the most closely linked gene. ***P* < 0.01.

Table 1. Summary of results from enrichment analysis of $\Delta AF > 0.8$ SNPs located in conserved noncoding elements. One significantly enriched term was chosen from each group of significantly enriched intercorrelated terms. Full lists of enriched terms and intercorrelations are presented in database S3, and the most enriched intercorrelated terms are presented in fig. S5. *P* values are Bonferroni-corrected. O/R, number of distinct nonoverlapping 1-Mb windows observed (O) and the average number of 1-Mb windows observed in 1000 random (R) samplings of the same number of genes (rounded to the nearest integer). TS, Thieler stage.

Database entry	Enriched term	Number of genes	*P*	Enrichment	Distinct loci (O/R)
		Gene Ontology biological process			
GO:0007399	Nervous system development	191	3.7×10^{-10}	1.7	154/155
GO:0045595	Regulation of cell differentiation	107	4.5×10^{-6}	1.8	94/91
GO:0045935	Positive regulation of nucleobase-containing compound metabolic process	122	2.0×10^{-5}	1.7	101/100
GO:0045165	Cell fate commitment	36	5.5×10^{-5}	2.9	31/32
GO:0007389	Pattern specification process	57	1.4×10^{-4}	2.2	43/44
GO:0009887	Organ morphogenesis	85	2.0×10^{-3}	1.8	72/73
GO:0048646	Anatomical structure formation involved in morphogenesis	75	2.8×10^{-3}	1.8	65/64
GO:0045892	Negative regulation of transcription, DNA-dependent	82	1.4×10^{-2}	1.7	62/62
GO:0034332	Adherens junction organization	13	1.5×10^{-2}	4.7	11/11
		Mouse Genome Informatics gene expression			
11853	TS23 diencephalon, lateral wall, mantle layer	109	3.9×10^{-25}	3.3	86/85
12449	TS23 medulla oblongata, lateral wall, basal plate, mantle layer	115	2.6×10^{-13}	2.3	90/89
2257	TS17 sensory organ	113	3.4×10^{-13}	2.3	98/99
1324	TS15 future brain	72	8.5×10^{-9}	2.4	61/61
		Mouse Genome Informatics phenotype			
MP:0010832	Lethality during fetal growth through weaning	240	7.5×10^{-17}	1.8	197/189
MP:0003632	Abnormal nervous system morphology	237	1.2×10^{-13}	1.7	191/193
MP:0005388	Respiratory system phenotype	127	1.7×10^{-7}	1.8	101/102
MP:0000428	Abnormal craniofacial morphology	109	1.4×10^{-6}	1.9	93/92
MP:0002925	Abnormal cardiovascular development	88	3.3×10^{-5}	1.9	73/73
MP:0005377	Hearing/vestibular/ear phenotype	73	1.8×10^{-4}	2.0	61/62

Cell fate determination was a strongly enriched GO category (enrichment factor = 4.9) (database S3) for genes near variants with high ΔAF. We examined the functional importance of 12 *SOX2*, 4 *KLF4*, and 1 *PAX2* high ΔAF SNPs associated with this GO category and where all 17 SNPs were distinct to domestic rabbits compared with other sequenced mammals. Electrophoretic mobility shift assay (EMSA) with nuclear extracts from mouse embryonic stem cell–derived neural stem cells revealed specific DNA-protein interactions (Fig. 3, fig. S7, table S7). Four probes, all from the *SOX2* region, showed a gel shift difference between wild and domestic alleles. Nuclear extracts from a mouse P19 embryonic carcinoma cell line before and after neuronal differentiation recapitulated these four gel shifts and revealed three additional probes, one in *PAX2* and two more in *SOX2*, that showed gel shift differences between wild-type and mutant probes only after neuronal differentiation. Thus, altered DNA-protein interactions were identified for 7 of the 17 high ΔAF SNPs that we tested, qualifying them as candidate causal SNPs that may have contributed to rabbit domestication.

Our results show that very few loci have gone to complete fixation in domestic rabbits and none at coding sites or at noncoding conserved sites. However, allele frequency shifts were detected at many loci spread across the genome, and the

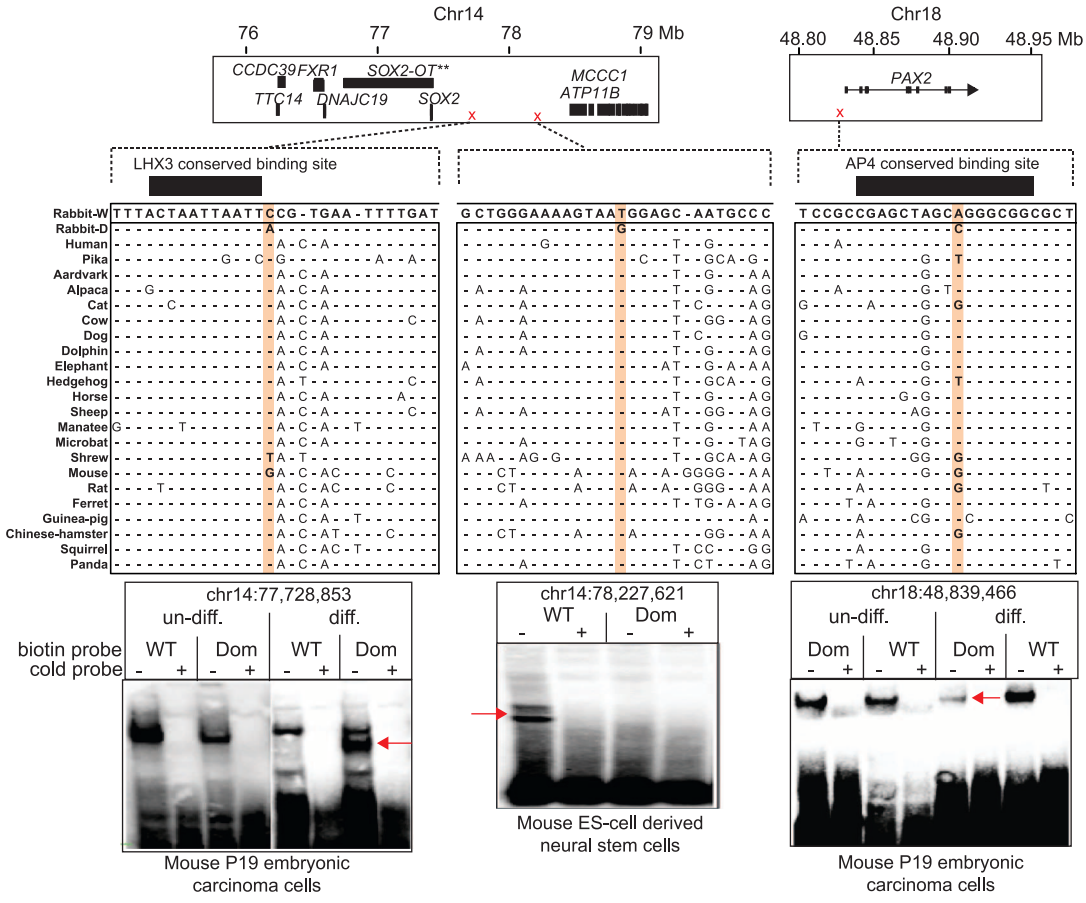

Fig. 3. Bioinformatic and functional analysis of candidate causal mutations. Three examples of SNPs near *SOX2* and *PAX2* where the domestic allele differs from other mammals. The locations of these three SNPs assessed with EMSA are indicated by red crosses on top. EMSA with nuclear extracts from embryonic stem cell–derived neural stem cells or from mouse P19 embryonic carcinoma cells before (un-diff) or after neuronal differentiation (diff) are shown for three SNPs. Exact nucleotide positions of polymorphic sites are indicated. Allele-specific gel shifts are indicated by arrows. WT, wild-type allele; Dom, domestic, the most common allele in domestic rabbits. Cold probes at 100-fold excess were used to verify specific DNA-protein interactions.

great majority of domestic alleles were also found in wild rabbits, implying that directional selection events associated with rabbit domestication are consistent with polygenic and soft sweep modes of selection (*18*) that primarily acted on standing genetic variation in regulatory regions of the genome. This stands in contrast with breed-specific traits in many domesticated animals that often show a simple genetic basis with complete fixation of causative alleles (*19*). Our finding that many genes affecting brain and neuronal development have been targeted during rabbit domestication is fully consistent with the view that the most critical phenotypic changes during the initial steps of animal domestication probably involved behavioral traits that allowed animals to tolerate humans and the environment humans offered. On the basis of these observations, we propose that the reason for the paucity of specific fixed domestication genes in animals is that no single genetic change is either necessary or sufficient for domestication. Because of the complex genetic background for tame behavior, we propose that domestic animals evolved by means of many mutations of small effect, rather than by critical changes at only a few domestication loci.

REFERENCES AND NOTES

1. C. Darwin, *On the Origins of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life* (John Murray, London, 1859).
2. J. A. Clutton-Brock, *Natural History of Domesticated Mammals* (Cambridge Univ. Press, Cambridge, 1999).
3. M. Carneiro *et al.*, *Mol. Biol. Evol.* **28**, 1801–1816 (2011).
4. Materials and methods are available as supplementary materials on *Science* Online.
5. M. Carneiro *et al.*, *Mol. Biol. Evol.* **29**, 1837–1849 (2012).
6. J. M. Smith, J. Haigh, *Genet. Res.* **23**, 23–35 (1974).
7. R. Nielsen *et al.*, *Genome Res.* **15**, 1566–1575 (2005).
8. K. Lindblad-Toh *et al.*, *Nature* **478**, 476–482 (2011).
9. M. M. Motazacker *et al.*, *Am. J. Hum. Genet.* **81**, 792–798 (2007).
10. K. Takahashi, S. Yamanaka, *Cell* **126**, 663–676 (2006).
11. C. Y. McLean *et al.*, *Nat. Biotechnol.* **28**, 495–501 (2010).
12. C.-J. Rubin *et al.*, *Nature* **464**, 587–591 (2010).
13. C.-J. Rubin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19529–19536 (2012).
14. M. V. Olson, *Am. J. Hum. Genet.* **64**, 18–23 (1999).
15. P. V. Tran *et al.*, *Nat. Genet.* **40**, 403–410 (2008).
16. K. Agger *et al.*, *Nature* **449**, 731–734 (2007).
17. H. Deng, K. Gao, J. Jankovic, *Nat. Rev. Neurol.* **8**, 203–213 (2012).
18. J. K. Pritchard, J. K. Pickrell, G. Coop, *Curr. Biol.* **20**, R208–R215 (2010).
19. L. Andersson, *Curr. Opin. Genet. Dev.* **23**, 295–301 (2013).

ACKNOWLEDGMENTS

This work was supported by grants from the National Human Genome Research Institute (U54 HG003067 to E.S.L.), European Research Council project BATESON to L.A., the Wellcome Trust (grants WT095908 and WT098051), the intramural research program of the NIH, NIAID (R.G.M.), the European Molecular Biology Laboratory, Programa Operacional Potencial Humano–Quadro de Referência Estratégica Nacional funds from the European Social Fund and Portuguese Ministério da Ciência, Tecnologia e Ensino Superior [postdoc grants to M.C. (SFRH/BPD/72343/2010) and R.C. (SFRH/BPD/64365/2009) and Ph.D. grant to J.M.A. (SFRH/BD/72381/2010)], a NSF international postdoctoral fellowship to J.M.G. (OISE-0754461), FEDER funds through the COMPETE program and Portuguese national funds through the Fundação para a Ciência e a Tecnologia (PTDC/CVT/122943/2010, PTDC/BIA-EVF/115069/2009, PTDC/BIA-BDE/72304/2006, and PTDC/BIA-BDE/72277/2006), the projects "Genomics and Evolutionary Biology" and "Genomics Applied to Genetic Resources" cofinanced by North Portugal Regional Operational Programme 2007/2013 (ON.2 – O Novo Norte) under the National Strategic Reference Framework and the European Regional Development Fund, travel grants to M.C. (COST Action TD1101), and Higher Education Commission in Pakistan (support for Sh.S.). We are grateful to L. Gaffney for assistance with figure preparation, P. C. Alves and S. Mills for providing the snowshoe hare sample, and S. Pääbo for hosting M.C., S.A., and R.C. Sequencing was performed by the Broad Institute Genomics Platform. Computer resources were supplied by BITS and UPPNEX at Science for Life Laboratory. The *O. cuniculus* genome assembly has been deposited in GenBank under the accession number AAGW02000000. The RNA sequencing data have been deposited in GenBank under the bioproject PRJNA78323, the rabbit genome resequencing data under the bioproject PRJNA242290, and the sequence capture data under the bioproject PRJNA221358. **Author contributions:** K.L.-T., F.D.P., and E.S.L. oversaw genome sequencing, assembly, and annotation performed by J.A., J.T.-M., J.J., D.H., J.L.C., and Sa.Y. B.A., D.B., M.R., and St.S. did Ensembl annotations. C.R.-G., V.D., L.F., R.G.M., and Z.P. contributed to the genome project. L.A., M.C., C.-J.R., N.F., and K.L.-T. led the domestication study, and A.M.B., N.R., and Sh.S. contributed with bioinformatic analyses. Sh.Y. and G.P. performed EMSA, and A.X. and K.F.-N. developed neural stem cells for EMSA. M.C., F.W.A., J.M.G., S.A., J.M.A., G.B., S.B., H.A.B., R.C., H.G., G.Q., and R.V. designed, performed, and analyzed the sequence capture experiment. L.A., M.C., C.-J.R., K.L.-T., N.F., and F.D.P. wrote the paper with input from other authors.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6200/1074/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S7
References (*20*–*61*)
Databases S1 to S5

21 March 2014; accepted 11 July 2014
10.1126/science.1253714