

An Expressed *Fgf4* Retrogene Is Associated with Breed-Defining Chondrodysplasia in Domestic Dogs

Heidi G. Parker,¹ Bridgett M. VonHoldt,² Pascale Quignon,¹ Elliott H. Margulies,³ Stephanie Shao,¹ Dana S. Mosher,¹ Tyrone C. Spady,¹ Abdel Elkhoulou,¹ Michele Cargill,^{4*} Paul G. Jones,⁵ Cheryl L. Maslen,⁶ Gregory M. Acland,^{7,8} Nathan B. Sutter,⁸ Keiichi Kuroki,⁹ Carlos D. Bustamante,¹⁰ Robert K. Wayne,² Elaine A. Ostrander^{1†}

Retrotransposition of processed mRNAs is a common source of novel sequence acquired during the evolution of genomes. Although the vast majority of retroposed gene copies, or retrogenes, rapidly accumulate debilitating mutations that disrupt the reading frame, a small percentage become new genes that encode functional proteins. By using a multibreed association analysis in the domestic dog, we demonstrate that expression of a recently acquired retrogene encoding fibroblast growth factor 4 (*fgf4*) is strongly associated with chondrodysplasia, a short-legged phenotype that defines at least 19 dog breeds including dachshund, corgi, and basset hound. These results illustrate the important role of a single evolutionary event in constraining and directing phenotypic diversity in the domestic dog.

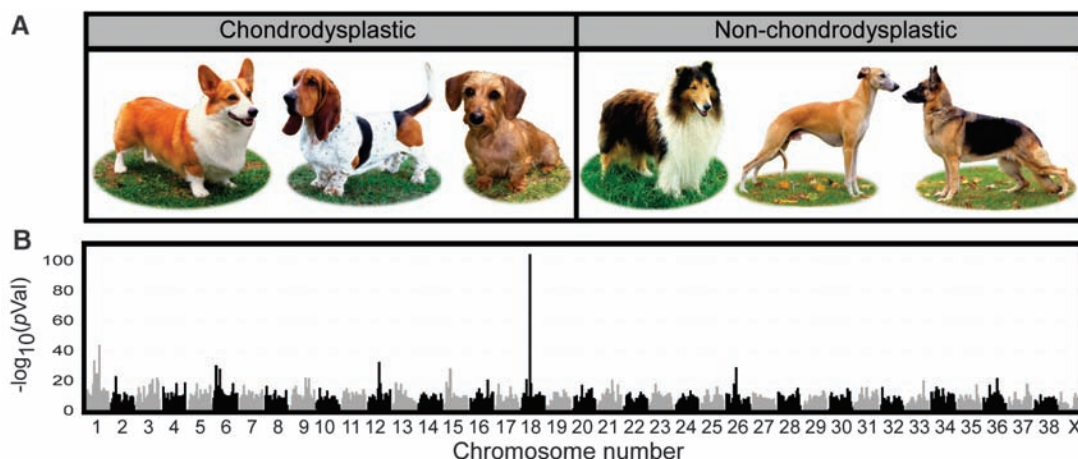
The domestic dog is arguably the most morphologically diverse species of mammal, and theories abound regarding the source of its extreme variation (1). Two such

¹Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA. ²Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA. ³Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA. ⁴Affymetrix Corporation, 3420 Central Expressway, Santa Clara, CA 95051, USA. ⁵The WALTHAM Centre for Pet Nutrition, Waltham on the Wolds, Leicestershire LE14 4RT, UK. ⁶Division of Cardiovascular Medicine, Oregon Health and Science University, Portland, OR 97239, USA. ⁷Baker Institute for Animal Health, Cornell University, Ithaca, NY 14853, USA. ⁸College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA. ⁹Comparative Orthopaedic Laboratory, University of Missouri, Columbia, MO 65211, USA. ¹⁰Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY, USA.

*Present address: Genetics Navigenics, Foster City, CA 94404, USA.

†To whom correspondence should be addressed. E-mail: eostrand@mail.nih.gov

Fig. 1. Results of a whole-genome association analysis for chondrodysplasia across 72 breeds of dog. **(A)** Examples of breeds used as cases (Pembroke Welsh corgi, basset hound, and dachshund pictured) and controls (collie, whippet, and German shepherd dog) in this analysis. [Photos credit: Mary Bloom, AKC] **(B)** Alternating shades of gray and black designate the chromosomal boundaries. The two highest peaks are found on chromosome 18 at bases 23,298,242 and 23,729,786 in CanFam2 assembly (16). The peaks are less than 0.5 Mb apart and appear merged in the graph.



theories rely on the structure and instability of the canine genome, either in an excess of rapidly mutating microsatellites (2) or an abundance of overactive short interspersed nuclear elements (SINEs) (3), to create increased variability from which to select for new traits. Another theory suggests that domestication has allowed for the buildup of mildly deleterious mutations that, when combined, create the variation observed in the domestic dog (4). The notion of gene duplication as a major cause of morphologic diversity has received little attention.

The majority of phenotypic variation in domestic dogs is found among, rather than within, the over 350 recognized domestic dog breeds. One aspect of interbreed variation is leg length, with some of the most striking short-legged breeds displaying limb morphology characteristic of chondrodysplasia, also known as short-limbed or disproportional dwarfism (table S1). The trait is a primary requirement in the American Kennel Club (AKC) breed standard for over a dozen domestic breeds including

dachshund, Pekingese, and basset hound, where it was found to be dominant and allelic on the basis of arranged crosses (5). The phenotype primarily affects the length of the long bones, with growth plates calcifying early in development, thus producing shortened bones with a curved appearance (Fig. 1A) (6, 7).

To identify the genetic foundations of breed-defining phenotypes such as canine chondrodysplasia, we developed a multibreed approach for mapping fixed canine traits. A total of 835 dogs from 76 distinct breeds that provided maximal coverage of phenotypic variation were genotyped by using the Affymetrix version 2.0 single-nucleotide polymorphism (SNP) chip (8, 9). Chondrodysplastic breeds, or cases, were defined on the basis of specific morphologic criteria set forth in each breed standard (8, 10) and comprised 95 dogs from eight breeds. The control or nonchondrodysplastic group included 702 dogs from 64 breeds lacking the above features (Fig. 1A and table S1).

Single-marker analysis revealed a strong association [odds ratio (OR) = 33.54] between a SNP on chromosome 18 (CFA18) at base position 23,298,242 (CanFam2) and the chondrodysplasia phenotype ($\chi^2 = 437$; $P = 9 \times 10^{-104}$ uncorrected; Fig. 1B). The second best peak of association was found at position 23,729,786; 431 kb telomeric to the first, with a P value of 2×10^{-57} . Because the P values are inflated because of population structure (4% of P values were less than 10^{-7}), we also performed independent Mann-Whitney U tests on the distribution of allele frequencies within the chondrodysplastic and control breeds. The two SNPs on CFA18 retained the strongest association with P values of 1.15×10^{-5} and 2.74×10^{-5} , respectively. The best haplotype across the chromosome spanned the five SNPs beginning at position 23,298,242 and ending at position 23,729,786 (uncorrected P value = 1.9×10^{-111}) (table S1).

Because registered members of a breed are expected to meet specific morphologic criteria, we hypothesized that breed-defining traits such as chondrodysplasia would be under strong

selective pressure. We compared heterozygosity in 139 cases and 173 controls genotyped at an additional 64 SNPs that spanned the associated region (table S2) and observed 125 kb (23,320,831 to 23,445,875) in which the cases displayed considerably lower amounts of heterozygosity than the controls did, indicative of a selective sweep (case average = 1.9%, control = 19.6%, $P = 6 \times 10^{-6}$, paired t test) (11–14).

We sequenced 54 amplicons in 44 dogs from 20 breeds (9 case and 11 control) with the goal of (i) identifying additional SNPs, (ii) identifying causative mutations, and (iii) finding the smallest haplotype shared among chondrodysplastic breeds (table S3). Of the 123 SNPs we identified, 50 formed a single continuous homozygous haplotype in all 26 chondrodysplastic dogs tested, covering about 24 kb (23,422,559 to 23,446,056) (Fig. 2A). A portion of the 3' untranslated region (3'UTR) of semaphorin 3c (*sema3c*), a putative thioredoxin domain containing one (*txndc1*) pseudogene, and two evolutionarily conserved sequences are contained within the shared haplotype (Fig. 2B).

An insert of about 5 kb starting at position 23,431,136 (fig. S1) was found by tiling polymerase chain reaction (PCR) amplicons across the homozygous region. This insert was present in all dogs from the original eight breeds and 11 of 12 additional breeds that fit at least two of the three chondrodysplastic criteria (175 dogs from 19 breeds) (8). Seven of the 175 short-legged dogs were heterozygous for the insert (table S4). The insert was not found in 204 medium- to long-legged dogs from 41 breeds that do not display the trait (table S4).

Although the insertion was unambiguously associated with chondrodysplasia, the initial analysis did not address whether the position of the insert or its specific content was causative. We therefore sequenced the insert with use of an Illumina Genome Analyzer (Illumina, Incorporated, San Diego, California). A library was first created from a gel-extracted long-range PCR product that spanned the entire insert from two unrelated chondrodysplastic dogs (dachshund and Scottish terrier). The sequence data were assembled by using Velvet algorithms (15). BLAT analysis (16) revealed a single contig with complete alignment at 100% identity to fibroblast growth factor 4 (*FGF4*), which is located on CFA18 at position 51,439,516; about 30 Mb from the insert.

With use of Sanger sequencing with primers designed from the annotated *FGF4* gene sequence together with the sequence surrounding the insertion site (table S5), we demonstrated that the insert contained a conserved *fgf4* retrogene. Neither the introns nor the upstream promoter sequences of the gene were present in the insert; however, all exons were present, with no alterations in the coding sequence, as well as the 3'UTR and polyadenylate [poly(A)] tail characteristic of retrotransposition of processed mRNA (Fig. 3).

Fig. 2. Observed heterozygosity in chondrodysplastic (red) and nonchondrodysplastic (black) breeds within the associated region on chromosome 18. (A) Graph of observed heterozygosity (H_o) across a 34-kb region on CFA18. Each point is the average H_o at one marker across all individuals within the group. The x axis shows the position on chromosome 18. The lines, red for chondrodysplastic and black for nonchondrodysplastic, show the trend in heterozygosity across the region by LOWESS (locally weighted least squares) best fit to the data. The average H_o for controls across the 24-kb homozygous region is 0.10. (B) Schematic of the region that is homozygous and identical in chondrodysplastic breeds. Gene 1 is a pseudogene similar to thioredoxin domain containing 1 (*txndc1*). Gene 2 is the 3' end of semaphorin 3c (*sema3c*). The green boxes labeled putative regulatory regions are conserved in both sequence and context in all mammals for which genome data are available. A 5-kb insertion (red rectangle) was found within the fourth LINE between the two putative regulatory elements. The insertion contains an *fgf4* retrogene. Arrangement of genes and conserved regions are per the CanFam2 assembly (16).

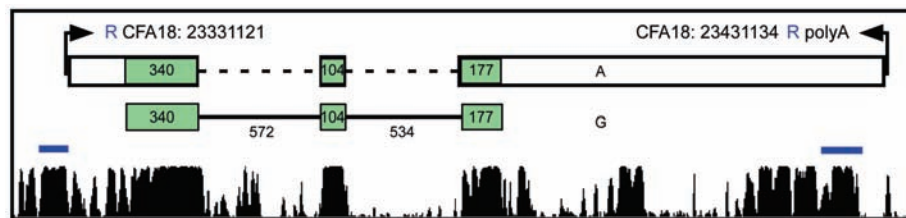
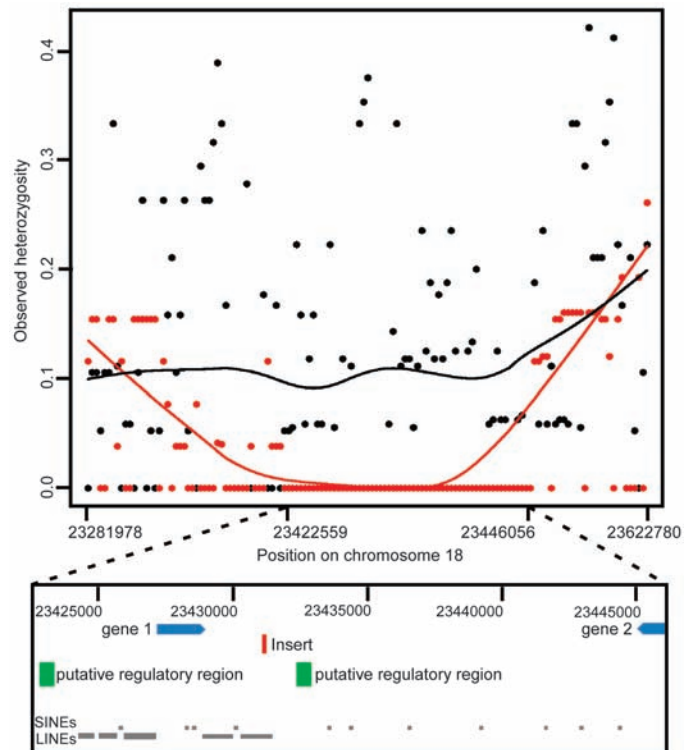


Fig. 3. Comparison of insert to source *FGF4* gene. The first row displays the alignment of the insert sequence to the source *FGF4* sequence. *FGF4* has three coding exons represented by the green boxes on the graph; begins at CFA18 position 51,439,420; and ends at position 51,441,146. All three exons are present in the insert, which aligns between positions 51,439,178 and 51,442,902. The insert includes 242 bases upstream of the start site and 1756 bases downstream of the stop codon followed by a poly(A) repeat. A 13-base sequence (AAGTCAGACAGAG) derived from the insert site, indicated by a blue R on the figure, is repeated at both ends of the insert. The second line shows the coding sequence of *FGF4* with the size of the exons and introns labeled. Alignment of the mouse promoter and enhancer sequences are indicated by the blue lines directly above the dog, human, mouse, and rat conservation track shown at the bottom of the figure (16) (Transcriptional Regulatory Element Database, <http://rulai.cshl.edu/cgi-bin/TRED/tred.cgi?process=home>) (35). Coding sequence is predicted on the basis of sequence similarity of translated proteins (accession no. XM_540801, www.ncbi.nlm.nih.gov).

To determine whether the retrogene was expressed, we searched for retrogene-specific sequences in complete cDNA of chondrodysplastic dogs. A single base at a position syntenic to chr18:51441601, 455 base pairs (bp) distal to the coding sequence of *FGF4*, differed between the retrogene and the source gene, with the former displaying an A nucleotide and the latter a

G, in all samples tested. Both A and G alleles were detected in cDNA generated from articular cartilage of the long bones of chondrodysplastic dogs (Fig. 4A), whereas only the G allele was detected in cDNA and genomic DNA samples from nonchondrodysplastic dogs (Fig. 4B).

Gene duplication through retrotransposition differs from a tandem duplication that may

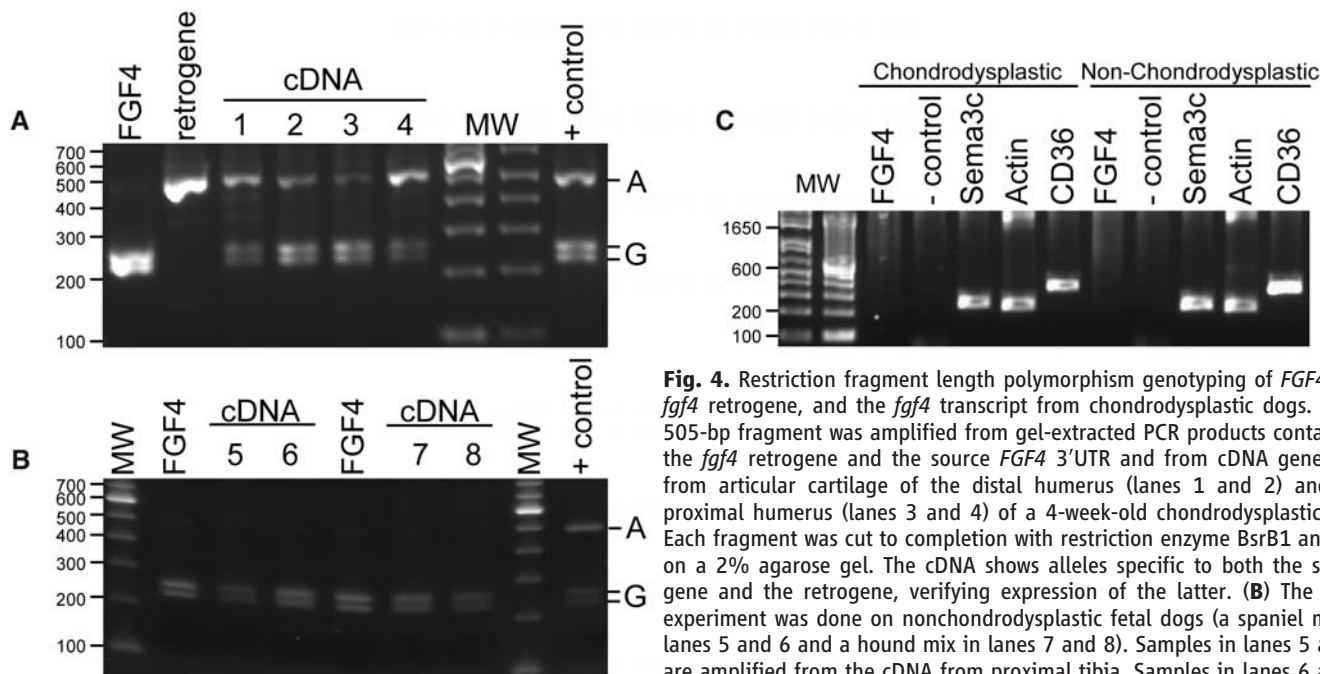


Fig. 4. Restriction fragment length polymorphism genotyping of *FGF4*, the *fgf4* retrogene, and the *fgf4* transcript from chondrodysplastic dogs. **(A)** A 505-bp fragment was amplified from gel-extracted PCR products containing the *fgf4* retrogene and the source *FGF4* 3'UTR and from cDNA generated from articular cartilage of the distal humerus (lanes 1 and 2) and the proximal humerus (lanes 3 and 4) of a 4-week-old chondrodysplastic dog. Each fragment was cut to completion with restriction enzyme BsrB1 and run on a 2% agarose gel. The cDNA shows alleles specific to both the source gene and the retrogene, verifying expression of the latter. **(B)** The same experiment was done on nonchondrodysplastic fetal dogs (a spaniel mix in lanes 5 and 6 and a hound mix in lanes 7 and 8). Samples in lanes 5 and 7 are amplified from the cDNA from proximal tibia. Samples in lanes 6 and 8 are from cDNA from distal femur. Genotypes from the source gene and cDNA are identical because no other copy of *FGF4* is present. **(C)** Genes were amplified in cDNA from articular cartilage from the proximal humerus in an adult chondrodysplastic (shih tzu) and a nonchondrodysplastic (Siberian husky) dog. Although RNA amounts were low in these tissues, expression of *CD36* and *Sema3c* was readily detected. However, expression of neither the source *FGF4* nor the *fgf4* retrogene could be detected.

simply double the gene dosage (17) because the retrogene must acquire a new promoter, likely with a different expression profile, in order to be active. To accomplish this, retrogenes often “borrow” contextual regulatory elements (18). We therefore assessed the expression of thrombospondin receptor (*CD36*) and *Sema3c* genes, which are upstream and downstream of the insert. A PCR-based assay on cDNA from the articular cartilage of fetal and neonatal dogs revealed expression of both genes in the growing limb (fig. S2). Further examination of expression in cartilage tissues from adult dogs shows that, although the surrounding genes were expressed, neither the source *FGF4* gene nor the *fgf4* retrogene were still expressed (Fig. 3C). The finding that the retrogene neither follows the expression pattern of neighboring genes nor is ubiquitously expressed implies that it has a specific time-sensitive role. The retrogene is inserted in the middle of a long interspersed nuclear element (LINE) with both LINEs and SINEs upstream (Fig. 2B). These transposable elements likely provide the regulatory machinery necessary to promote expression of the *fgf4* retrogene (19) with localization and temporal control coming from the intact 3'UTR (20).

We hypothesize that atypical expression of the *FGF4* transcript in the chondrocytes causes inappropriate activation of one or more of the fibroblast growth factor receptors such as *FGFR3*. An activating mutation in *FGFR3* is responsible for >95% of achondroplasia cases, the most common form of dwarfism in humans, and 60 to 65% of hypochondroplasia cases, a

human syndrome that is more similar in appearance to breed-defining chondrodysplasia [reviewed in (21)]. *FGF4* induces the expression of sprouty genes, which interfere with the ubiquitin-mediated degradation of the FGF receptors including *FGFR3*, and overexpression of the sprouty genes can cause chondrodysplastic phenotypes in both mice and humans (22, 23).

The chondrodysplastic breeds were developed in many different countries for a variety of occupations (10). On the basis of genomic analysis of population structure, they do not share a recent common ancestry (24, 25). However, because we find a common haplotype of 24 kb surrounding the *fgf4* retrogene in 19 short-legged breeds, it is likely the chondrodysplastic phenotype arose only once, before the division of early dogs into modern breeds. Thereafter, the retrogene and its associated phenotype were both maintained and propagated by breeders for purposes specific to each breed.

To further explore the origin of the *fgf4* retrogene, we compared haplotypes from the source gene, the retrogene, and the insertion site in both dogs and their wild progenitor, the gray wolf. The ancestor of all chondrodysplastic breeds would have needed to carry both a source gene with the rare haplotype found in the retrogene and the 24-kb haplotype that defines the insertion site (fig. S3 and table S6). This combination was not found in any of the dogs that we tested but was identified in wolves from Europe and the Middle East, supporting fossil evidence that these populations contributed to the early development of the dog (26, 27).

Although retrogenes are recognized as an important source of novel functional elements found between recently diverged species (28–30), little is known about the relation between retrotransposition and phenotypic variation within species (30, 31). We have found a single retrotransposition event producing a conserved, expressed retrogene that has strongly focused the evolutionary direction of morphological change in the dog because at least 12% of American breeds share a common phenotype and the retrogene. This retrogene is actively segregating within the species, has a coding sequence that is identical to that of the source gene, and to the best of our knowledge is the only example of a functional retrogene found in morphologically distinct populations of a single species that is actively maintained by selection. If such rare mutational events or “sports,” as Charles Darwin referred to them in *The Origin of Species* (32), happen only in the evolution of domestic animals, then these systems may be less informative for understanding the origin of evolutionary novelty in wild species. However, if the molecular phenomenon we have observed represents a class of genomic change associated with dramatic phenotypic evolution, such as that characteristic of adaptive radiation (18, 33, 34), then such genetic changes might be keystone molecular innovations.

References and Notes

1. R. K. Wayne, E. A. Ostrander, *Trends Genet.* **23**, 557 (2007).
2. J. W. Fondon 3rd, H. R. Garner, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 18058 (2004).
3. E. F. Kirkness *et al.*, *Science* **301**, 1898 (2003).

4. F. Cruz, C. Vila, M. T. Webster, *Mol. Biol. Evol.* **25**, 2331 (2008).
5. C. R. Stockard, *The Genetic and Endocrinic Basis for Differences in Form and Behavior* (Wistar Institute of Anatomy and Biology, Philadelphia, 1941).
6. S. Heuertz et al., *Eur. J. Hum. Genet.* **14**, 1240 (2006).
7. N. Palmer, in *Pathology of Domestic Animals*, K. V. F. Jubb, P. C. Kennedy, N. Palmer, Eds. (Academic Press, San Diego, CA, 1993), vol. 1.
8. Materials and methods are available as supporting material on Science Online.
9. C. Drogemuller et al., *Science* **321**, 1462 (2008).
10. American Kennel Club, *The Complete Dog Book: Official Publication of the American Kennel Club* (Howell Book House, New York, ed. 19, 1998).
11. K. L. Simonsen, G. A. Churchill, C. F. Aquadro, *Genetics* **141**, 413 (1995).
12. C. Colby, S. M. Williams, *Genetics* **140**, 1129 (1995).
13. N. B. Sutter et al., *Science* **316**, 112 (2007).
14. J. P. Pollinger et al., *Genome Res.* **15**, 1809 (2005).
15. D. R. Zerbino, E. Birney, *Genome Res.* **18**, 821 (2008).
16. <http://genome.ucsc.edu>
17. N. H. C. Salmon Hillbertz et al., *Nat. Genet.* **39**, 1318 (2007).
18. N. Vinckenbosch, I. Dupanloup, H. Kaessmann, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3220 (2006).
19. C. Feschotte, *Nat. Rev. Genet.* **9**, 397 (2008).
20. D. Fraidraich, R. Lang, C. Basilio, *Dev. Biol.* **204**, 197 (1998).
21. W. A. Horton, G. P. Lunstrum, *Rev. Endocr. Metab. Disord.* **3**, 381 (2002).
22. G. Minowada et al., *Development* **126**, 4465 (1999).
23. C. Guo et al., *Cell. Signal.* **20**, 1471 (2008).
24. H. G. Parker et al., *Genome Res.* **17**, 1562 (2007).
25. H. G. Parker et al., *Science* **304**, 1160 (2004).
26. M. Germonpre et al., *J. Archaeol. Sci.* **36**, 473 (2009).
27. M. V. Sablin, G. A. Khlopachev, *Curr. Anthropol.* **43**, 795 (2002).
28. J. Brosius, *Genetica* **107**, 209 (1999).
29. A. C. Marques, I. Dupanloup, N. Vinckenbosch, A. Reymond, H. Kaessmann, *PLoS Biol.* **3**, e357 (2005).
30. H. Kaessmann, N. Vinckenbosch, M. Long, *Nat. Rev. Genet.* **10**, 19 (2009).
31. S. Ohno, *Semin. Cell Dev. Biol.* **10**, 517 (1999).
32. C. R. Darwin, *The Origin of Species by Means of Natural Selection* (John Murray, London, 1872).
33. E. Betran, Y. Bai, M. Motiwale, *Mol. Biol. Evol.* **23**, 2191 (2006).
34. J. Rohozinski, C. E. Bishop, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11695 (2004).
35. T. Nowling, C. Bernadt, L. Johnson, M. Desler, A. Rizzino, *J. Biol. Chem.* **278**, 13696 (2003).
36. We thank D. Babcock and C. Deginn for technical assistance, J. L. Cook for help with tissue identification and acquisition, L. Niswander for thoughtful discussions, and E. Giniger for careful reading of the manuscript. We gratefully acknowledge the dog owners who provided samples, the American Kennel Club–Canine Health Foundation, Affymetrix Corporation, and the Intramural Program of the National Human Genome Research Institute. Funded by NSF grants 0733033 (R.K.W.) and 516310 (C.D.B.) and NIH grants 5R01EY006855 and 1R24GM082910 (G.M.A.) and 1R01GM83606 (C.D.B.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1173275/DC1

Materials and Methods

Figs. S1 to S3

Tables S1 to S6

References

9 March 2009; accepted 24 June 2009

Published online 16 July 2009;

10.1126/science.1173275

Include this information when citing this paper.

Loss of Function of a Proline-Containing Protein Confers Durable Disease Resistance in Rice

Shuichi Fukuoka,^{1*} Norikuni Saka,² Hironori Koga,³ Kazuko Ono,¹ Takehiko Shimizu,⁴ Kaworu Ebana,¹ Nagao Hayashi,⁵ Akira Takahashi,⁵ Hirohiko Hirochika,⁶ Kazutoshi Okuno,⁷ Masahiro Yano¹

Blast disease is a devastating fungal disease of rice, one of the world's staple foods. Race-specific resistance to blast disease has usually not been durable. Here, we report the cloning of a previously unknown type of gene that confers non-race-specific resistance and its successful use in breeding. *Pi21* encodes a proline-rich protein that includes a putative heavy metal-binding domain and putative protein-protein interaction motifs. Wild-type *Pi21* appears to slow the plant's defense responses, which may support optimization of defense mechanisms. Deletions in its proline-rich motif inhibit this slowing. *Pi21* is separable from a closely linked gene conferring poor flavor. The resistant *pi21* allele, which is found in some strains of *japonica* rice, could improve blast resistance of rice worldwide.

The use of resistance genes (*R* genes) that confer race-specific resistance is a cost-effective strategy for the control of disease in crops. *R* genes are a key component of disease

resistance and often are associated with a hypersensitive response (HR), according to the gene-for-gene concept (1). Most *R* genes cloned in plants contain conserved gene structures (2), and mutation of the pathogen effectors that trigger *R* gene-mediated resistance leads to a loss of resistance, suggesting a substantial vulnerability of this defense mechanism (3). In contrast, resistance controlled by quantitative trait loci (QTLs) is usually non-race-specific (4), and its durability might be a consequence of decreased selective pressure for pathogens to overcome host resistance. Despite the importance of resistance QTLs in crops, their molecular basis remains largely unknown, except *Lr34* and *Yr36*, which house genes that differ structurally from previously reported *R* genes (5, 6).

The resistance to blast disease caused by *Magnaporthe oryzae* conferred by QTLs in rice

(*Oryza sativa* L.) lacks HR, yet it restricts the development of lesions (7). Some cultivars carrying resistance QTLs, such as the *Pi21* locus (8), have maintained resistance throughout a century of cultivation. However, the reintroduction of resistance and undesirable agricultural traits, including grain characteristics from donors, has prevented the use of potential genetic resources for the development of elite cultivars for the past 80 years (9). The causes of the troublesome association have long been debated and could include either tight linkage of genes that control independent traits, known as "linkage drag," or pleiotropic effects of the target gene on other traits (10). Such factors pose penalties on crop production, yet research on the cost to agricultural crops is limited (10). To identify the molecular basis of *Pi21*-mediated resistance and to settle the long-term debate about the cause of associated undesirable characteristics, we undertook map-based cloning of *pi21*, a recessive allele conferring resistance to rice blast.

Leaves of AA-*pi21*, a near-isogenic line (NIL) carrying *pi21* in the genetic background of a susceptible cultivar, had smaller lesions than did the susceptible cultivar (AA) in both a field evaluation and a greenhouse inoculation test (Fig. 1, A and B, and figs. S1A and S2). The resistant *pi21* allele had a consistent effect against all 10 of the widely distributed races of *M. oryzae* studied, although the resistance was incomplete as compared with that triggered by *R* genes (Fig. 1C). Histological examination showed that cells infected with a virulent race were still alive and intact in both AA and AA-*pi21* 40 hours after inoculation (Fig. 1D and table S1), in contrast to those infected with an avirulent race, which induces *R* gene-dependent hypersensitive cell death (table S1). However, cytoplasmic granules, the first sign of cell death, were frequently observed in AA-*pi21* 96 hours after inoculation (Fig. 1E).

¹QTL Genomics Research Center, National Institute of Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan. ²Mountainous Region Agricultural Research Institute, Aichi Agricultural Research Center, Inahasi, Toyota, Aichi 441-2513, Japan. ³Bioproduction Sciences, Ishikawa Prefectural University, Suematsu 1-308, Nonoichi-machi, Ishikawa 921-8836, Japan. ⁴Institute of the Society for Techno-Innovation of Agriculture, Forestry and Fisheries, Ippaizuka, Tsukuba, Ibaraki 305-0854, Japan. ⁵Division of Plant Sciences, National Institute of Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan. ⁶Division of Genome and Biodiversity Research, National Institute of Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan. ⁷Graduate School of Life and Environmental Sciences, University of Tsukuba, Tennohdai 1-1-1, Tsukuba, Ibaraki 305-8572, Japan.

*To whom correspondence should be addressed. E-mail: fukusan@affrc.go.jp

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

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of June 27, 2015):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/325/5943/995.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2009/07/16/1173275.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/325/5943/995.full.html#related>

This article **cites 29 articles**, 16 of which can be accessed free:

<http://www.sciencemag.org/content/325/5943/995.full.html#ref-list-1>

This article has been **cited by** 7 article(s) on the ISI Web of Science

This article has been **cited by** 17 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/325/5943/995.full.html#related-urls>

This article appears in the following **subject collections**:

Genetics

<http://www.sciencemag.org/cgi/collection/genetics>