Positional Cloning of the Human Quantitative Trait Locus Underlying Taste Sensitivity to Phenylthiocarbamide

Un-kyung Kim *et al.*

Science **299**, 1221 (2003);
DOI: 10.1126/science.1080190

5. P. V. Tobias, *Olduvai Gorge, Vol. 4a and 4b: The Skulls, Endocasts and Teeth of* Homo habilis (Cambridge Univ. Press, Cambridge, 1991).
6. B. Wood, *Koobi Fora Research Project Vol. 4, Hominid Cranial Remains* (Clarendon, Oxford, UK, 1991).
7. G. H. Curtis, R. L. Hay, in *Calibration of Hominid Evolution*, W. W. Bishop, J. A. Miller, Eds. (Scottish Academic, Edinburgh, 1972), pp. 289–302.
8. R. C. Walter, P. C. Manega, R. L. Hay, R. E. Drake, G. H. Curtis, *Nature* **354**, 145 (1991).
9. R. C. Walter, P. C. Manega, R. L. Hay, *Quat. Int.* **13/14**, 37 (1992).
10. W. A. Berggren, D. V. Kent, C. C. Swisher III, M.-P. Aubry, *S.E.P.M. (Soc. Sediment. Geol.) Spec. Publ. No. 54* (1995), p. 129.
11. C. B. Stringer, in *Major Topics in Primate and Human Evolution*, B. Wood, L. Martin, P. Andrews, Eds. (Cambridge Univ. Press, Cambridge, 1986), pp. 266–294.
12. G. P. Rightmire, *Am. J. Phys. Anthropol.* **90**, 1 (1993).
13. C. P. Groves, *A Theory of Human and Primate Evolution* (Clarendon, Oxford, UK, 1989).
14. B. Wood, *Nature* **355**, 783 (1992).
15. _____, in *Species, Species Concepts, and Primate Evolution*, W. H. Kimbel, L. B. Martin, Eds. (Plenum, New York, 1993), pp. 485–522.
16. _____, *J. Hum. Evol.* **36**, 115 (1999).
17. L. S. B. Leakey, M. D. Leakey, *Nature* **202**, 5 (1964).
18. P. V. Tobias, *Nature* **209**, 953 (1966).
19. A. Walker, in Homo erectus: *Papers in Honor of Davidson Black*, B. A. Sigmon, J. S. Cybulski, Eds. (Univ. Toronto Press, Toronto, Canada, 1981), pp. 193–215.
20. D. C. Johanson *et al., Nature* **327**, 205 (1987).
21. R. F. McMahon, in *The Mollusca: Vol. 6: Ecology*, W. D. Russell-Hunter, Ed. (Academic, New York, 1983), pp. 505–561.
22. J. A. Lee-Thorpe, N. J. van der Merwe, *S. Afr. J. Sci.* **83**, 712 (1987).
23. R. J. Blumenschine, C. R. Peters, *J. Hum. Evol.* **34**, 565 (1998).
24. R. J. Blumenschine, *J. Hum. Evol.* **29**, 21 (1995).
25. S. D. Capaldo, *J. Hum. Evol.* **33**, 555 (1997).
26. H. T. Bunn, E. M. Kroll, *Curr. Anthropol.* **27**, 431 (1986).
27. C. R. Peters, R. J. Blumenschine, *J. Hum. Evol.* **29**, 321 (1995).
28. _____, in *Four Million Years of Hominid Evolution in Africa: Papers in Honour of Dr. Mary Douglas Leakey's Outstanding Contribution in Paleoanthropology*, C. Magori, C. B. Saanane, F. Schrenk, Eds. *(Kaupia*, Darmstadter Beitrage zur Naturgeschichte, Darmstadt, Germany, 1996), pp. 175–221.
29. We thank the Tanzanian government for permission to conduct fieldwork at Olduvai; C. Kilembe, O. Kileo, C. Msuya, F. Ndunguru, G. K. Olle Moita, and J. Pareso for facilitating this work; T. D. White for comments on the hominid taxonomic interpretations; C. S. Feibel for identifying *Corbicula* and providing the reference on its environmental tolerances; S. C. Anton for comments on table S2; and P. V. Tobias for comments on the submitted manuscript. S. R. Copeland, G. F. Mollel, J. Temba, G. Peter, and A. Venance provided excavation and mapping assistance. A. Venance and A. E. Cushing found OH 65 during excavation. The research was funded by NSF (grants SBR-9000099, SBR-9601065, and BCS-0109027), Rutgers University, Wenner-Gren Foundation for Anthropological Research, National Geographic Society, L. S. B. Leakey Foundation, Smithsonian Institution, Harvard University, University of Georgia, and Boise Fund.

Supporting Online Material
www.sciencemag.org/cgi/content/full/299/5610/1217/DC1
SOM Text
Fig. S1
Tables S1 to S3

24 June 2002; accepted 13 December 2002

Positional Cloning of the Human Quantitative Trait Locus Underlying Taste Sensitivity to Phenylthiocarbamide

Un-kyung Kim,[1] Eric Jorgenson,[2] Hilary Coon,[3,4] Mark Leppert,[3] Neil Risch,[2,5] Dennis Drayna[1]*

The ability to taste the substance phenylthiocarbamide (PTC) has been widely used for genetic and anthropological studies, but genetic studies have produced conflicting results and demonstrated complex inheritance for this trait. We have identified a small region on chromosome 7q that shows strong linkage disequilibrium between single-nucleotide polymorphism (SNP) markers and PTC taste sensitivity in unrelated subjects. This region contains a single gene that encodes a member of the TAS2R bitter taste receptor family. We identified three coding SNPs giving rise to five haplotypes in this gene worldwide. These haplotypes completely explain the bimodal distribution of PTC taste sensitivity, thus accounting for the inheritance of the classically defined taste insensitivity and for 55 to 85% of the variance in PTC sensitivity. Distinct phenotypes were associated with specific haplotypes, which demonstrates that this gene has a direct influence on PTC taste sensitivity and that sequence variants at different sites interact with each other within the encoded gene product.

The inability to taste PTC (*1, 2*) was long believed to be a simple Mendelian recessive trait (*3–8*). Over time, however, many reports emerged that contradicted this model (*9–11*). Linkage studies have been equally conflicting. Initial studies provided strong support for linkage to the KEL blood group antigen (later determined to reside on chromosome 7q3) (*12, 13*), but other studies failed to provide significant support for this linkage (*14*). The only genome-wide linkage survey was performed with the related compound propyl thiouracil. This study produced evidence for linkage to loci on chromosome 5p and a suggestion of linkage to markers on chromosome 7q31, at a distance of ~35 cM from KEL (*15*).

We performed a genome-wide linkage analysis with the Utah Centre d'Etude du Polymorphisme Humain (CEPH) families (*16–18*), using a blind sorting test to measure each individual's PTC sensitivity thresholds (*18–20*). We demonstrated strong support for a major locus on chromosome 7q, close to KEL (*21, 22*), with a critical region spanning about 4 Mb in the region of D7S661, with a maximum lod score (logarithm of odds ratio for linkage) of 8.85 (*22*).

Bioinformatic analyses (*18*) indicated that the ~4-Mb region on chromosome 7q (Fig. 1A) contains more than 150 genes, including the one that encodes the KEL blood group antigen, which confirms previous linkage studies (*12, 13*). In addition, this region contains a number of TAS2R bitter taste receptor genes (*23*) and odorant receptor–like (OR-like) genes (*24*). We evaluated all TAS2Rs (nine genes) and OR-like genes (seven genes) (table S1) as candidates by sequencing the entire single coding exon, the 3′ untranslated region, and 300 base pairs (bp) upstream in individuals within families showing linkage to chromosome 7q; we observed numerous sequence variants (*25, 26*). One of these variants demonstrated strong association with taste phenotype across different CEPH families (χ^2 analysis, $P < 10^{-10}$), suggesting that it may be the functional change or close to the functional change(s). To more fully understand linkage and linkage disequilibrium (LD) relationships in this region, we performed further analysis by means of 50 single-nucleotide polymorphisms (SNPs) (table S2) at an average spacing of 50 kb across this interval. These SNPs revealed crossover breakpoints in the Utah CEPH families that reduced the minimal region to 2.6 Mb (fig. S1).

With these 50 SNPs, we observed strong LD between taster status and markers in only

[1]National Institute on Deafness and Other Communication Disorders, National Institutes of Health, 5 Research Court, Rockville, MD 20850, USA. [2]Department of Genetics, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, USA. [3]Department of Human Genetics, University of Utah Medical Center, 15 North 2030 East, Salt Lake City, UT 84112, USA. [4]Department of Psychiatry, University of Utah Medical Center, Red Butte Health Center, 546 Chipeta Way, Salt Lake City, UT 84112, USA. [5]Division of Research, Kaiser Permanente, Oakland, CA 94612, USA.

*To whom correspondence should be addressed. E-mail: drayna@nidcd.nih.gov

one portion of this 2.6-Mb interval (Fig. 1B). This was observed initially in the chromosome 7–linked families (12 families containing 107 individuals) and subsequently in unrelated nontasters from both the CEPH sample (an additional 8 individuals) and in a second replication population [the National Institutes of Health (NIH) sample, 15 nontaster and 14 taster Caucasians, 7 nontaster and 9 taster East Asians]. We observed significant LD across a 150-kb region, extending from about 139,835,000 to 139,981,000 bp on the chromosome 7 genomic sequence (*27*). In the NIH sample of 45 individuals, χ^2 analysis (equivalent to r^2) and delta statistics showed clear peak values for each measure within the bacterial artificial chromosome (BAC) RP11-707F14 (AC073647.9) ($P <$

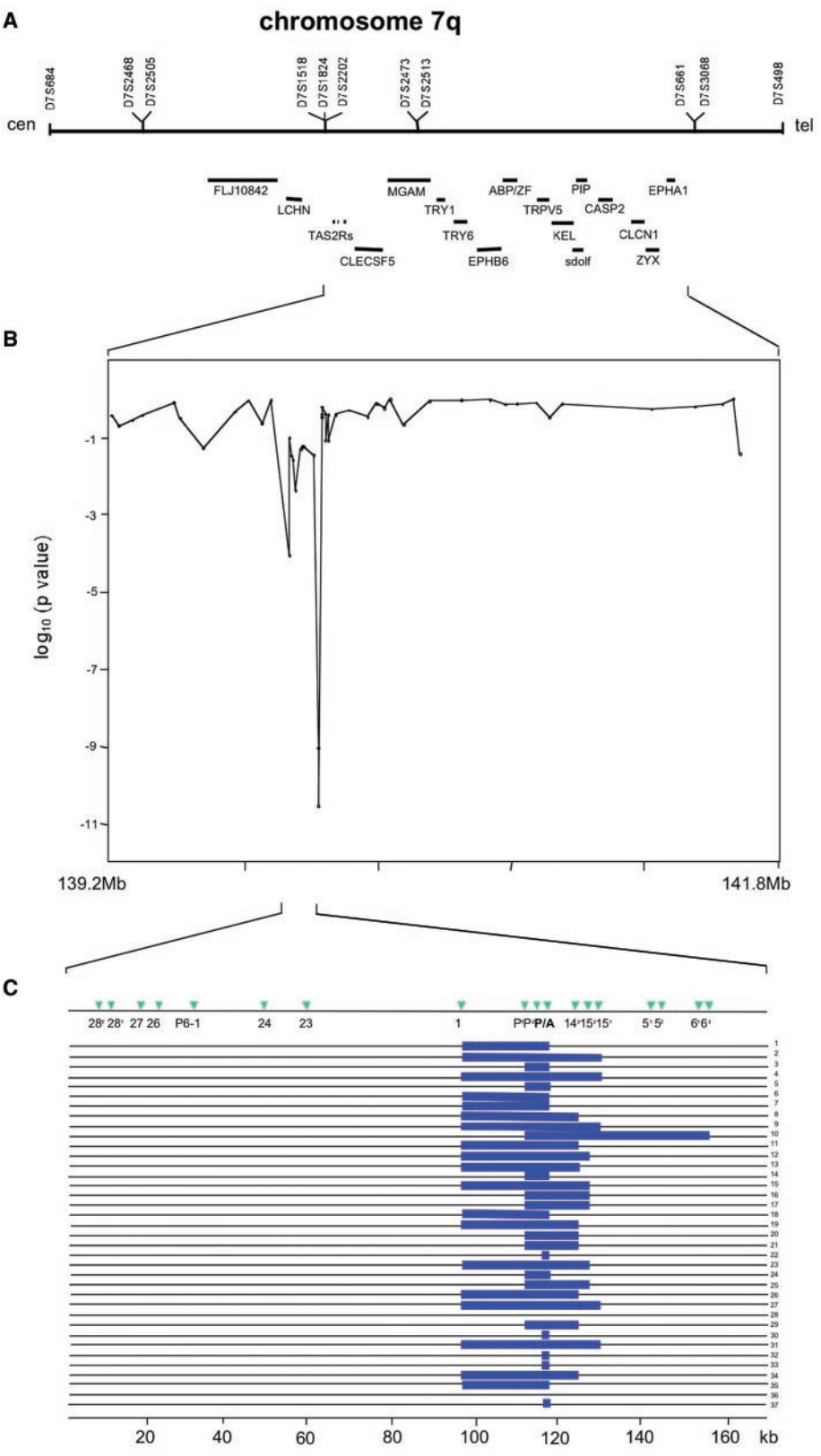

Fig. 1. Genetic localization of the *PTC* gene. (**A**) Genetic mapping of the *PTC* locus. Linkage analysis using existing CEPH genotypes gave a gene interval bounded by D7S684 on the centromeric side and D7S661 on the telomeric side (~4 Mb). Additional SNP genotyping in the Utah families reduced the interval on the centromeric side to marker D7S1518/2202. The known genes and their respective positions in this interval are indicated. Physical distance is colinear with the scale in (B). (**B**) Evaluation of SNP LD was performed with χ^2 tests using Yates correction for continuity, with 45 unrelated individuals from the NIH sample (*x* axis, physical distance on the chromosome 7q genomic DNA sequence; *y* axis, *P* values of χ^2 comparison of allele frequencies in tasters with those in nontasters). Each gradation marks 500 kb. (**C**) Comparison of shared haplotypes in the 37 nontaster individuals in the combined Utah/NIH sample. Individuals are listed in order of PTC sensitivity, with those showing poorest PTC sensitivity at the top and proceeding to more sensitive individuals at the bottom. Colored bars mark the extent of the conserved nontaster haplotype in each individual. Name and physical location of each SNP are shown at top, and physical distance in kb is shown at bottom. Individuals 28 and 36 are random individuals with no demonstrated linkage to chromosome 7.

10^{-10}) (Fig. 1B; figs. S2 and S3) at identical locations in the Caucasian and East Asian subgroups as well as for the Mantel-Haenszel combined χ^2 (table S3). In a group of 37 unrelated nontaster individuals (12 Utah individuals and 25 individuals from the NIH sample who collectively had the poorest PTC sensitivities) (*18*), the physical distance over which these individuals carried unambiguous haplotypes sharing the same SNP alleles extended an average of 61 kb, with the minimal shared region extending from 42,445 to 72,141 bp in this BAC, a distance of 29,696 bp (Fig. 1C). Bioinformatic and gene prediction analyses (*18*) revealed that the only gene in this 29.7-kb interval was the TAS2R bitter receptor gene in which we originally identified strong LD.

This gene, which we have designated *PTC*, consists of 1002 bp in a single exon, encoding a seven–transmembrane domain, heterotrimeric guanine nucleotide–binding protein (G protein)–coupled receptor that shows 30% amino acid identity with human TAS2R7, the most closely related member of this family. Within this gene, we identified three common SNPs, all of which result in amino acid changes in the protein (Table 1). The $Ala^{49} \rightarrow$ Pro (A49P) variant demonstrated a strong association overall with taster status in the Utah sample and an even stronger association in the NIH replication sample (Table 2). The association of taster status with the Val^{262} allele was similarly strong in both the Utah and NIH samples (Table 2). To better understand the effect of these SNPs, we investigated the haplotypes in this gene (*18*).

Table 1. Polymorphisms within the *PTC* gene.

Position		Allele	Frequency	Amino acid encoded
Base pair	Amino acid			
145	49	C	0.36	Pro
		G	0.64	Ala
785	262	C	0.38	Ala
		T	0.62	Val
886	296	G	0.38	Val
		A	0.62	Ile

Haplotype analysis in the Utah and NIH samples revealed two predominant haplotypes at the three SNPs in this gene. Named in the order of the three SNPs [A49P, V262A ($Val^{262} \rightarrow$ Ala), and I296V ($Ile^{296} \rightarrow$ Val)], the nontaster haplotype AVI and taster haplotype PAV accounted for 47% and 49% of all haplotypes, respectively, in the European sample and for 30% and 70%, respectively, in the East Asian sample. Europeans also possessed the presumed recombinant taster haplotype AAV at a frequency of 3%. The haplotype association with taster status was more definitive than for individual SNPs; the strongest association with nontaster status is for the AVI homozygote, followed by the compound heterozygote AVI/AAV (Table 3).

Because of the broad and continuous distribution of PTC sensitivity in the population, we went on to analyze PTC scores as a quantitative trait (*18*). There was a consistent and significant difference in PTC scores between diplotypes in both the Utah and the NIH samples, consistent across racial groups (table S4). PAV homozygotes had the highest mean PTC scores (Utah, 10.69; NIH, 10.00), and PAV heterozygotes had slightly but significantly lower mean PTC scores (Utah, 9.65; NIH, 8.81) than the PAV homozygotes (Utah sample, $\chi^2 = 8.41$, $P = 0.0037$; NIH replication sample, $t = 3.29$, $P = 0.0017$). AVI homozygotes had the lowest mean PTC scores (Utah, 4.31; NIH, 1.86). Thus, the taster PAV form of the gene displays a heterozygote effect, with two copies conferring greater PTC sensitivity than a single copy. The difference in mean PTC score between the rare AAV/AVI heterozygotes and the AVI homozygotes was significant in the NIH sample ($t = 5.44$, $P = 5.41 \times 10^{-5}$) and tended toward significance in the Utah family sample ($\chi^2 = 2.39$, $P = 0.122$). PAV/AAV heterozygotes were not significantly different from PAV/AVI heterozygotes ($\chi^2 = 0.58$, $P = 0.45$).

Differences in PTC score by diplotype (that is, genotypes at multiple variable sites with consideration of haplotype) in the Utah families were also highly significant in a multivariate analysis ($\chi^2 = 148.95$, $P < 10^{-33}$) (*18*). Sex and the haplotype effect explain 59.7% of the total variance in PTC scores. Analysis of variance of the NIH sample confirmed these results ($F = 152.73$, $P < 10^{-32}$), with 84.8% of the variance explained by the haplotype effect. The differences were also significant in both the Caucasian subgroup of the replication sample ($F = 78.60$, $P < 10^{-18}$) and the East Asian subgroup ($F = 139.02$, $P < 10^{-11}$).

The bimodal distribution of PTC scores is a combination of the underlying distributions of the *PTC* diplotypes (Fig. 2). The appearance of bimodality is driven by the distribution of the common AVI homozygote, PAV/AVI heterozygote, and PAV homozygote diplotypes. The mode of inheritance of PTC taste sensitivity has been a subject of controversy (*2, 10, 11*). To determine whether there was evidence for additional genetic contributions to PTC score, we examined the heritability in subsets of the Utah sample. In the subgroups that were large enough to give accurate estimates, heritability was 0.26 ± 0.19 (83 subjects in 20 families) in the PAV/AVI subgroup and 0.50 ± 0.33 in the AVI/AVI subgroup (46 subjects in 17 families). The increase in heritability in the loss-of-function diplotype group (AVI/AVI) indicates that there may be other genetic factors that interact with *PTC* and that can restore some measure of taste sensitivity in this group. For Caucasians and East Asians, our results are largely consistent with a model of a major recessive quantitative trait locus modified by either a polygenic (*10*) or a single locus (*11*) residual background effect.

Because of the high frequency of the PAV and AVI haplotypes in the population, we sought to determine which haplotype represents the original form of the *PTC* gene. We sequenced this gene in six primate species: humans and one individual each from chimpanzee, lowland gorilla, orangutan, crab-eating macaque (an Old World monkey), and black-handed spider monkey (a New World monkey), representing more than 25 million years of evolutionary divergence. All the nonhuman primates were homozygous for the PAV form, which indicates that the AVI form arose in

Table 2. Effect of homozygosity for SNPs on phenotype. The third SNP, I296V, was in complete LD with V262A (and thus gave identical results to V262A) except in one African-American subject.

Homozygous SNP	Sample	No. of subjects (total no.)		χ^2	*P* value
		Nontasters	Tasters		
Ala^{49}	Utah	48 (51)	21 (129)	27.23	1.81×10^{-7}
	NIH	22 (23)	3 (61)	72.74	1.61×10^{-16}
Val^{262}	Utah	38 (51)	14 (129)	23.40	1.10×10^{-6}
	NIH	21 (23)	0 (61)	74.44	6.83×10^{-17}

Table 3. Haplotype association with taster phenotypes.

Haplotypes	Sample	No. of subjects	
		Nontasters	Tasters
AVI/AVI	Utah	38	14
	NIH	21	0
AVI/AAV	Utah	10	7
	NIH	1	3
*/PAV	Utah	3	108
	NIH	1	58

*Any haplotype found in the sample. No AAV homozygotes were observed in either sample.

humans after they diverged from the nearest common primate ancestors.

Five different haplotypes were observed worldwide (table S5). In Europeans and Asians, the taster haplotype PAV and the nontaster haplotype AVI make up the vast majority of haplotypes present. Two additional haplotypes, PVI and AAI, were observed only in individuals of sub-Saharan African ancestry, which is consistent with other reports of increased gene haplotype diversity in this population (*28*). The common nontaster AVI haplotype was observed in all populations except Southwest Native Americans, who were exclusively homozygous for the PAV haplotype, consistent with the reported low frequency of nontasters in this population (*2*). Thus, overall, the worldwide distribution of these haplotypes is consistent with the large anthropologic literature on the distribution of this phenotype (*29, 30*).

The amino acid substitutions in the *PTC* protein may affect the function of this protein in several ways. Position 49 resides in the predicted first intracellular loop, and this SNP represents a major amino acid alteration—from proline in tasters to alanine in nontasters. The SNPs at position 262, in the predicted sixth transmembrane domain, and position 296, in the predicted seventh transmembrane domain, specify relatively conserved amino acid changes—alanine to valine and valine to isoleucine, respectively. On the basis of phenotype data (table S4), we hypothesize that the substitutions at positions 49 and 262 significantly alter the biochemical function of this protein, whereas the substitution at position 296 modifies the function more subtly. These alterations could affect coupling to its cognate G proteins on the intracellular side of the plasma membrane, as has been observed for other variants in the first intracellular loop (*31, 32*) or in other portions of these proteins (*33*). Given that PTC and other compounds that contain the N–C=S moiety are both bitter and toxic in large doses, it will be of interest to determine how the nontaster allele rose to such high frequency, especially in the European population.

Substantial variation in taste sensitivity exists in humans (*34*), and, given the great degree of sequence diversity and variation in bitter taste receptor genes (*35*), we hypothesize that much of this phenotypic variation is genetic in origin. Understanding the nature of this variation, especially variation in bitter taste, and its relationship to diet and other behaviors such as smoking may have important implications for human health (*36, 37*).

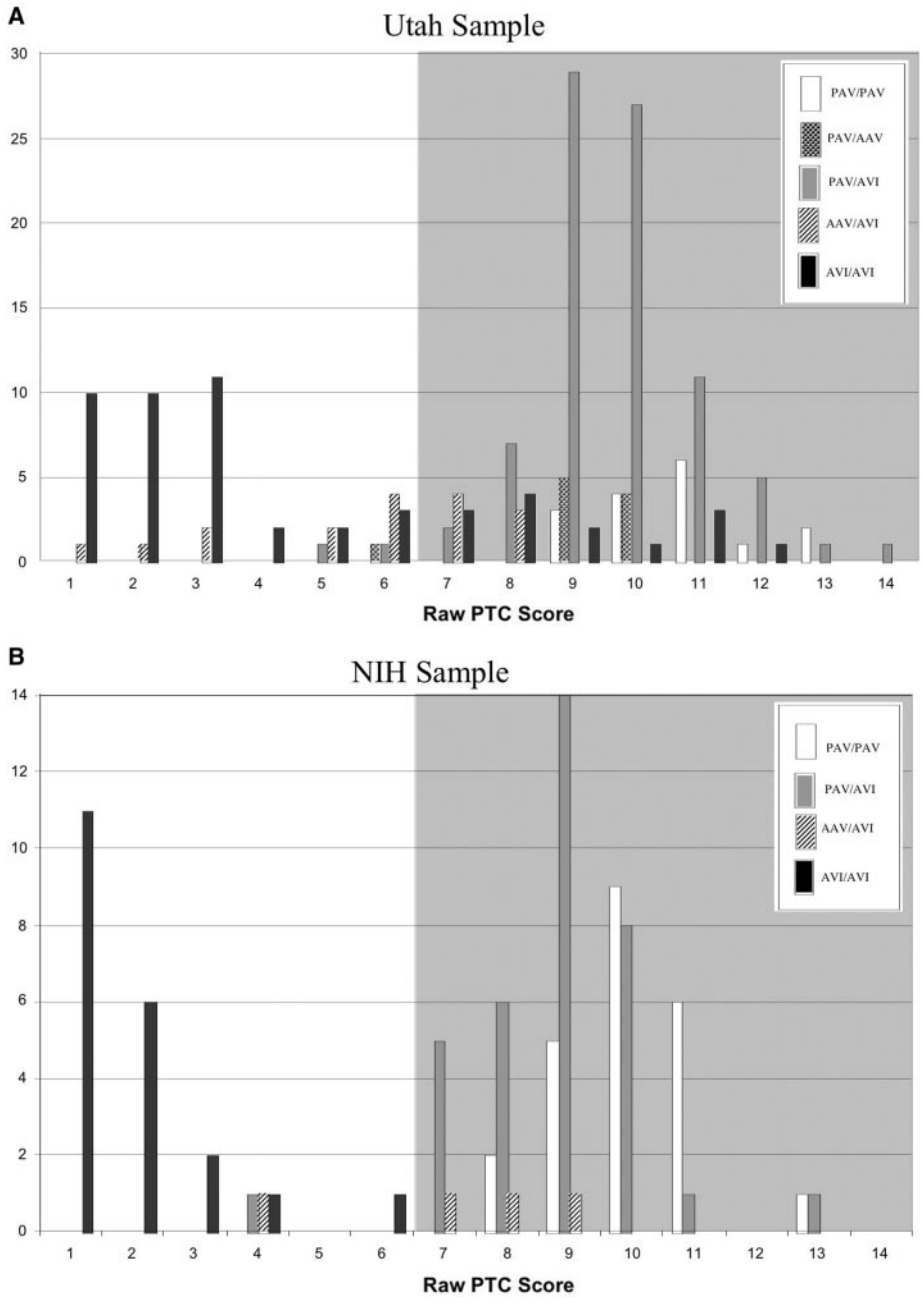

Fig. 2. Distribution of PTC scores for groups defined by PTC diplotype. (**A**) Utah family sample. (**B**) NIH sample of unrelated individuals. Scores in the shaded regions were classified as tasters and those in the unshaded regions were classified as nontasters in the dichotomous assignment of phenotype.

References and Notes

1. Anonymous, *Science* **73** (suppl.), 14 (1931).
2. S.-W. Guo, D. Reed, *Ann. Hum. Biol.* **28**, 111 (2001).
3. L. H. Snyder, *Science* **74**, 151 (1931).
4. S. G. Levit, G. V. Soboleva, *J. Genet.* **30**, 389 (1935).
5. A. F. Blakeslee, *Proc. Natl. Acad. Sci. U.S.A.* **18**, 120 (1932).
6. B. F. Lee, *Ohio J. Sci.* **34**, 337 (1934).
7. H. Harris, H. Kalmus, *Ann. Eugen. London* **15**, 24 (1949).
8. B. B. Merton, *Acta Genet. Statist. Med.* **8**, 114 (1958).
9. D. S. Falconer, *Ann. Eugen.* **13**, 211 (1946–47).
10. B. M. Reddy, D. C. Rao, *Genet. Epidemiol.* **6**, 413 (1989).
11. J. M. Olson, M. Boehnke, K. Neiswanger, A. F. Roche, R. M. Siervogel, *Genet. Epidemiol.* **6**, 423 (1989).
12. E. A. Chautard-Freire-Maia *et al.*, *Ann. Hum. Genet.* **38**, 191 (1974).
13. P. M. Conneally, M. Dumont-Driscoll, R. S. Huntzinger, W. E. Nance, C. E. Jackson, *Hum. Hered.* **26**, 267 (1976).
14. A. Spence *et al.*, *Hum. Genet.* **67**, 183 (1984).
15. D. Reed *et al.*, *Am. J. Hum. Genet.* **64**, 1478 (1999).
16. J. Dausset *et al.*, *Genomics* **6**, 575 (1990).
17. NIH/CEPH Collaborative Mapping Group, *Science* **258**, 67 (1992).
18. See supporting data on *Science* Online.
19. H. Harris, H. Kalmus, *Ann. Eugen. London* **15**, 24 (1949).
20. H. Kalmus, *Ann. Hum. Genet.* **22**, 222 (1958).
21. D. A. Prodi *et al.*, *Am. J. Hum. Genet.* **71** (suppl. 4), 464 (2002).
22. D. Drayna *et al.*, *Hum. Genet.*, in press.
23. E. Adler *et al.*, *Cell* **100**, 693 (2000).
24. L. Buck, R. Axel, *Cell* **65**, 175 (1991).
25. B. Ewing, L. Hillier, M. C. Wendl, P. Green, *Genome Res.* **8**,175 (1998).
26. D. Gordon, C. Abajian, P. Green, *Genome Res.* **8**, 195 (1998).
27. National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/genome/guide/human).

28. J. C. Stephens *et al*., *Science* **293**, 489 (2001).
29. W. C. Boyd, *Genetics & the Races of Man: An Introduction to Modern Physical Anthropology* (Little, Brown, Boston, 1950), pp. 276–294.
30. D. Tills, A. Kopec, R. Tills, *The Distribution of Human Blood Groups and Other Polymorphisms, Supplement* (Oxford University Press, Oxford, ed. 1, 1983).
31. C. Nabhan, Y. Xiong, L. Y. Xie, A. B. Abou-Samra, *Biochem. Biophys. Res. Commun.* **212**, 1015 (1995).
32. B. F. O'Dowd *et al*., *J. Biol. Chem.* **263**, 15985 (1988).
33. G protein receptor database: www.gpcr.org.
34. A. Blakeslee, T. Salmon, *Proc. Natl. Acad. Sci. U.S.A.* **21**, 84 (1935).
35. T. Ueda *et al*., *Biochem. Biophys. Res. Commun.* **285**, 147 (2001).
36. B. Tepper, *Am. J. Hum. Genet.* **63**, 1271 (1998).
37. M.-A. Enoch, C. Harris, D. Goldman, *Addict. Behav.* **26**, 399 (2001).
38. Supported by the Division of Intramural Research, National Institute on Deafness and Other Communication Disorders (grant Z01-000046-04); the National Human Genome Research Institute (grant T32 HG00044); the National Center for Research Resources (grant M01-RR00064) to the Huntsman General Clinical Research Center at the University of Utah; the W. M. Keck Foundation; and the George S. and Delores Doré Eccles Foundation. We are indebted to the members of the Utah CEPH families and the NIH research volunteers for their participation. We thank T. Friedman, A. Griffith, S. Sullivan, and the anonymous reviewers for helpful comments on the manuscript. All subjects provided informed consent, indicating the nature and possible consequences of this study, under University of Utah Medical Center protocol 6090-96 and National Institute of Neurological Diseases and Stroke NIH protocol DC-01-230.

Supporting Online Material
www.sciencemag.org/cgi/content/full/299/5610/1221/DC1
Materials and Methods
Figures S1 to S3
Tables S1 to S5

6 November 2002; accepted 21 January 2003

Impact of Genetic Manipulation on the Fitness of *Anopheles stephensi* Mosquitoes

Flaminia Catteruccia,[1] H. Charles J. Godfray,[1,2] Andrea Crisanti[1,3*]

Genetic modification of mosquitoes offers exciting possibilities for controlling malaria, but success will depend on how transformation affects the fitness of modified insects. The expression of an exogenous gene, the mutations caused by its insertion, and inbreeding while transformed lines are established can all lead to reductions in fitness. Factors influencing fitness were investigated in cage experiments with four lines of transgenic *Anopheles stephensi*, a vector species of human malaria. The results indicate direct costs of the introduced transgene in at least three out of the four lines, as well as an apparent cost of the inbreeding involved in making transgenic homozygotes.

The idea that malaria might be controlled by introducing genes into mosquito populations to hinder or block parasite transmission and hence to make them refractory was proposed more than three decades ago (*1*), but it is only in the past 2 years that it has become a possibility. Robust gene transfer technologies involving transposon-mediated transformation have been developed for *Anopheles stephensi* and *Anopheles gambiae*, the major malaria vectors in India and Africa, respectively (*2*, *3*), and mosquitoes have been engineered to express a gene that significantly reduces vector competence in an experimental malaria model (*4*). Several other control strategies have been suggested, all requiring the release of genetically modified individuals into the environment (*5*, *6*). However, before the chances of success of these control measures can be judged and any such release can be contemplated by regulatory authorities, it will be essential to understand the relative fitness and population ecology of genetically modified mosquitoes (*7*).

To begin to address this issue, we have studied the fate of introduced genes in populations of caged mosquitoes by using mixtures of transformed and wild-type mosquitoes. We worked with a series of *A. stephensi* transgenic lines containing either the enhanced green fluorescent protein (EGFP) or the red fluorescent protein (from the coral *Discosoma*, DsRed) as a phenotypic marker (*8*, *9*). The genes had been inserted into the mosquito chromosomes without a drive mechanism (e.g., an active transposon). Our null hypothesis is that the frequency of the introduced gene should be subject only to genetic drift, whereas any fitness costs of the gene or the process of transformation will lead to a reduction in frequency.

Four transformed mosquito lines (named IV, V_{D12}, AsML12, and MinRED1) (*2*, *10*) were used in this study. These lines do not carry "malaria-refractory" genes and therefore are not potential control targets. Nevertheless, they represent a predictive model for evaluating the fitness costs of transgenic technologies, as they were developed by using *minos*-based transposons carrying different combinations of genes and promoters in addition to EGFP or DsRed (table S1). Transgenic homozygous lines were propagated for more than 30 generations without detectable loss of the fluorescent marker or evidence of transposon mobilization (*11*). During this period, the transgenic lines did not show major differences in longevity, feeding rate, larval development time, or egg laying rate in comparison with wild-type mosquitoes.

We established a series of mosquito populations from equal numbers (100 first instar larvae) of wild-type mosquitoes from a long-established laboratory culture and homozygous individuals from the transgenic lines. For each generation, 200 larvae were chosen randomly out of a population of between 1000 and 2000 individuals, assayed for whether they were homozygous, heterozygous, or wild type for the EGFP or DsRed phenotype, and then allowed to emerge and mate. Four days later they were fed on mouse blood and, after a further 48 hours, allowed to oviposit to produce the next generation of mosquitoes. Visual examination under ultraviolet light allowed us unambiguously to distinguish wild-type larvae from those heterozygous or homozygous for fluorescent transgenes according to the intensity of the green or red fluorescence (Fig. 1) (*11*). Two experiments were run for lines IV and AsML12, and single experiments for V_{D12} and MinRED1.

In all populations of mosquitoes, the frequency of the transgenic allele in the population decreased sharply over a few generations (Fig. 2, A to D). Under the null hypothesis of no effects of the transgene or transformation on mosquito fitness, we would expect a Hardy-Weinberg equilibrium to be established after one generation, with 50% heterozygotes and equal numbers of homozygotes (*12*). In all cases, transgene frequencies were lower than expected in the first generation and tended to decline over time; they became extinct within 4 to 16 generations.

Lines IV and V_{D12} involved the same transposon construct, which was inserted at different sites (table S1). Analysis of the integration site in line IV showed that the transposon had disrupted a coding sequence (*2*) with similarities to the *Drosophila melanogaster chaoptin* precursor gene (*13*). This

[1]Department of Biological Sciences, SAF Building, Imperial College London, South Kensington Campus, London SW7 2AZ, UK. [2]NERC Centre for Population Biology, Imperial College London, Silwood Park Campus, Ascot, Berkshire SL5 7PY, UK. [3]University of Perugia, Via del Giochetto, 06100 Perugia, Italy.

*To whom correspondence should be addressed. Email: acrs@ic.ac.uk