

High-Frequency Persistence of an Impaired Allele of the Retroviral Defense Gene *TRIM5 α* in Humans

Sara L. Sawyer,¹ Lily I. Wu,² Joshua M. Akey,³ Michael Emerman,^{1,2,4} and Harmit S. Malik^{1,4,*}

¹Division of Basic Sciences

²Division of Human Biology

Fred Hutchinson Cancer Research Center

Seattle, Washington 98109

³Department of Genome Sciences

University of Washington

Seattle, Washington 98109

Summary

The intracellular *TRIM5 α* protein successfully inhibits HIV-1 infection in rhesus monkeys, but not in humans [1–6]. A few amino acids in the virus-interacting SPRY domain [7] were found to be responsible for most of this anti-viral specificity [8–10], raising the possibility that genetic variation among humans could result in *TRIM5 α* proteins with a spectrum of potencies. We found several nonsynonymous SNPs at the human *TRIM5* locus, but only one of these (H43Y) was found to have a significant functional consequence. We demonstrate that H43Y impairs *TRIM5 α* restriction of two distantly related retroviruses. H43Y lies in the RING domain of *TRIM5 α* and may negatively affect its putative E3 ubiquitin ligase activity. This detrimental allele dates back to before the African diaspora and is found at a frequency of 43% in indigenous Central and South Americans. We suggest that relaxed constraint due to a recent period of low retroviral challenge has allowed the deleterious H43Y mutation to persist and even to expand after the bottleneck that occurred upon human migration to the New World. The unexpectedly high frequency of an impaired retroviral restriction allele among humans is likely to have a significant impact on our ability to ward off future retroviral challenges.

Results and Discussion

To test the hypothesis that variation in the human gene pool encodes *TRIM5 α* proteins with a spectrum of antiviral potencies, we surveyed global variation in the human *TRIM5 α* gene. We sequenced 4879 bp of genomic sequence, from 2 bp upstream of the start codon in exon 2 to the stop codon in exon 8 (a large intron of 10.5 kb between exons 4 and 5 was not sequenced) from 37 geographically diverse, indigenous humans (74 chromosomes from Africa, the Middle East, Southeast Asia, Europe, and Central and South America, all selected from the Coriell Human Variation collection of the NIGMS repository). A total of 20 SNPs were identified (Figures 1A and 1C), which ranged in frequency from 1% to 50% (Figure 1D). Sequence from one chimpanzee and one gorilla

was obtained in order to assign ancestral (white, major allele) and derived (black, minor allele) states for each SNP (Figure 1A) and to determine changes that have been fixed specifically in the human lineage in the last 5 million years (Figure 1B). Positive selection of *TRIM5 α* throughout primate evolution has been concentrated in the SPRY domain [10, 11] and five replacement changes in the SPRY domain have been fixed in the human lineage since its common ancestor with chimpanzee (Figure 1B). Because the SPRY domain encodes determinants of capsid recognition [8–10, 12, 13], these changes may reflect new specificities that were fixed during the last 5 million years of human evolution.

We also find nucleotide polymorphisms throughout the length of the gene, including six nonsynonymous SNPs (nsSNPs). Bioinformatic analyses predict that three of these, H43Y, V112F, and R238W, could potentially affect the function of the *TRIM5 α* protein (see the Supplemental Data available with this article online). To explore this possibility, we assayed each of the six nsSNPs for any effect on *TRIM5 α* function in CRFK cells (feline renal fibroblasts), which have been used to test the activity of exogenously expressed primate proteins against retroviruses because they themselves have no intrinsic retroviral restriction [3]. CRFK cells were stably transduced with human *TRIM5 α* or variants carrying each individual derived nsSNP. These cell lines, expressing roughly equal levels of each *TRIM5 α* variant (Figure 2A), were challenged with HIV-1 expressing GFP and assayed for percent infection. As has been previously reported [1], “wild-type” human *TRIM5 α* has only a moderate ability to restrict HIV-1 replication in this assay (Figure 2B). Whereas V112F, R136Q, R238W, G249D, and H419Y had no significant effect on *TRIM5 α* restriction, we found that H43Y significantly reduced viral restriction to a level similar to that of CRFK cells expressing no exogenous *TRIM5 α* .

Positive selection of the *TRIM5 α* gene predates primate lentiviruses like HIV-1 by many millions of years [10] and was likely driven by interactions with older endogenous and exogenous retroviruses. Moreover, because HIV-1 was only recently introduced into the human population [14], standing genetic variation in human *TRIM5 α* is unlikely to have been shaped by HIV-1. In contrast to their restriction of HIV-1, human cells have been found to efficiently restrict N-MLV [15], a murine γ retrovirus that is more closely related to the endogenous retroviruses that have been active during primate evolution [16]. Human *TRIM5 α* has been shown to be necessary and sufficient for the N-MLV restriction [2–5]. Thus, restriction against N-MLV is likely to be a more evolutionarily relevant activity for *TRIM5 α* . When we tested the various nsSNPs for N-MLV restriction, we found that V112F, R136Q, R238W, G249D, and H419Y did not affect *TRIM5 α* restriction, but H43Y was again seriously compromised (Figure 2C). H43Y was found to have no effect on the related B-MLV (Figure 2D), another MLV variant that is resistant to *TRIM5 α* restriction [2–5, 15].

*Correspondence: hsmalik@fhcrc.org

⁴These authors contributed equally to this work.

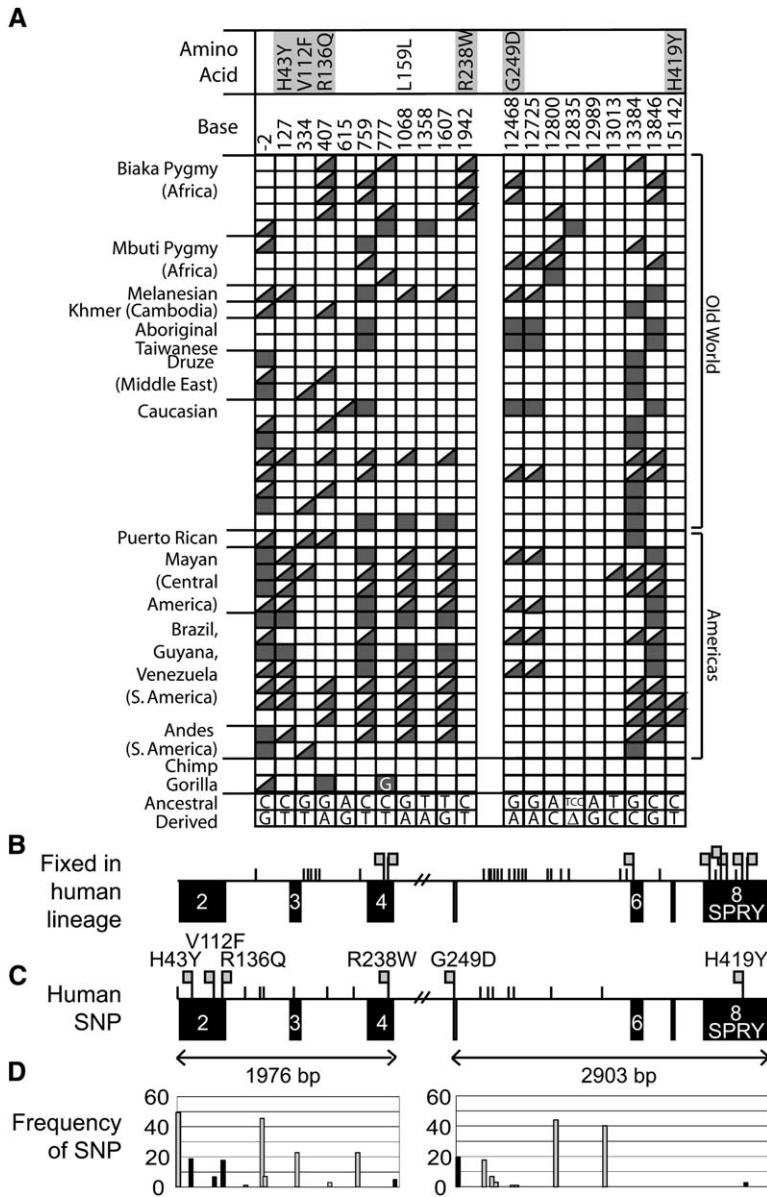


Figure 1. Polymorphism and Divergence in the Human *TRIM5 α* Gene

(A) The grid illustrates derived and ancestral states at SNP sites in the 37 humans sampled. Each row represents one human, and each column represents a SNP site (with base coordinates beginning at 1 = “A” of the ATG translational start site). White boxes indicate homozygous for the ancestral state (as defined by the major human allele and the chimpanzee sequence at the bottom), black boxes indicate homozygous for the derived state, and split boxes indicate heterozygosity. SNPs in exons are indicated along the top by their amino acid coordinates. We also obtained the sequence of one gorilla (bottom), which shares the human polymorphism at the -2 position and has a unique base at 777. The unsequenced intron falls between sites 1942 and 12468 and is indicated by a break in the grid.

(B and C) A schematic illustrates the *TRIM5 α* genomic structure with coding exons indicated by black boxes. The twenty SNPs that were identified are shown spatially on the gene diagram (C). The chimpanzee and gorilla sequences were used to define fixed human-specific changes (B). In both diagrams, flags indicate replacement changes, whereas stems indicate synonymous changes.

(D) The frequency of each SNP is plotted directly below its position in the schematic in (C) (light bars are silent SNPs, dark bars are nsSNPs).

To assay whether the impaired retroviral restriction seen with exogenously expressed *TRIM5 α* -H43Y translates to altered susceptibility in human cells, we tested B-lymphocytes from four individuals of our South American population set: one homozygous for H43 (“wild-type”), two homozygous for H43Y, and one heterozygous at this site. We infected these cell lines with N-MLV, B-MLV, or HIV-1. Human cells are profoundly resistant to infection by N-MLV due to endogenously expressed *TRIM5 α* [15]; correspondingly, B-lymphocytes homozygous for “wild-type” *TRIM5 α* were very poorly infected by N-MLV (Figure 3A). However, we found that cells from individuals homozygous for the H43Y mutation could be infected with N-MLV about 100-fold more efficiently whereas B-lymphocytes from heterozygous individuals had an intermediate phenotype (we confirmed that the heterozygote cell line transcribes both *TRIM5 α* alleles). Therefore, there was a direct correlation between the *TRIM5 α* genotype and the ability of human

cells to restrict this γ retrovirus. Although these cell lines likely differ at many other genetic loci, two lines of evidence suggest that the *TRIM5* locus is responsible for the decreased restriction of N-MLV. First, *TRIM5 α* has been previously shown to be the major human determinant of N-MLV restriction [2-5]. Second, we find that there is less than a 2-fold difference between the cell lines in their ability to be infected by B-MLV (Figure 3B), a variant of MLV that differs from N-MLV by a single amino acid change in its capsid that renders it immune to *TRIM5 α* restriction [2-5, 15]. These results and their congruence with what we observed by expressing exogenous *TRIM5 α* alleles in CRFK cells (Figure 2) demonstrate that genetic variation in the *TRIM5* locus has resulted in a loss of retroviral restriction in the human population. We also challenged the B-lymphocyte cells with HIV-1 and found that there was no significant effect of the H43Y change (Figure 3C). This is not surprising because previous studies have shown that although

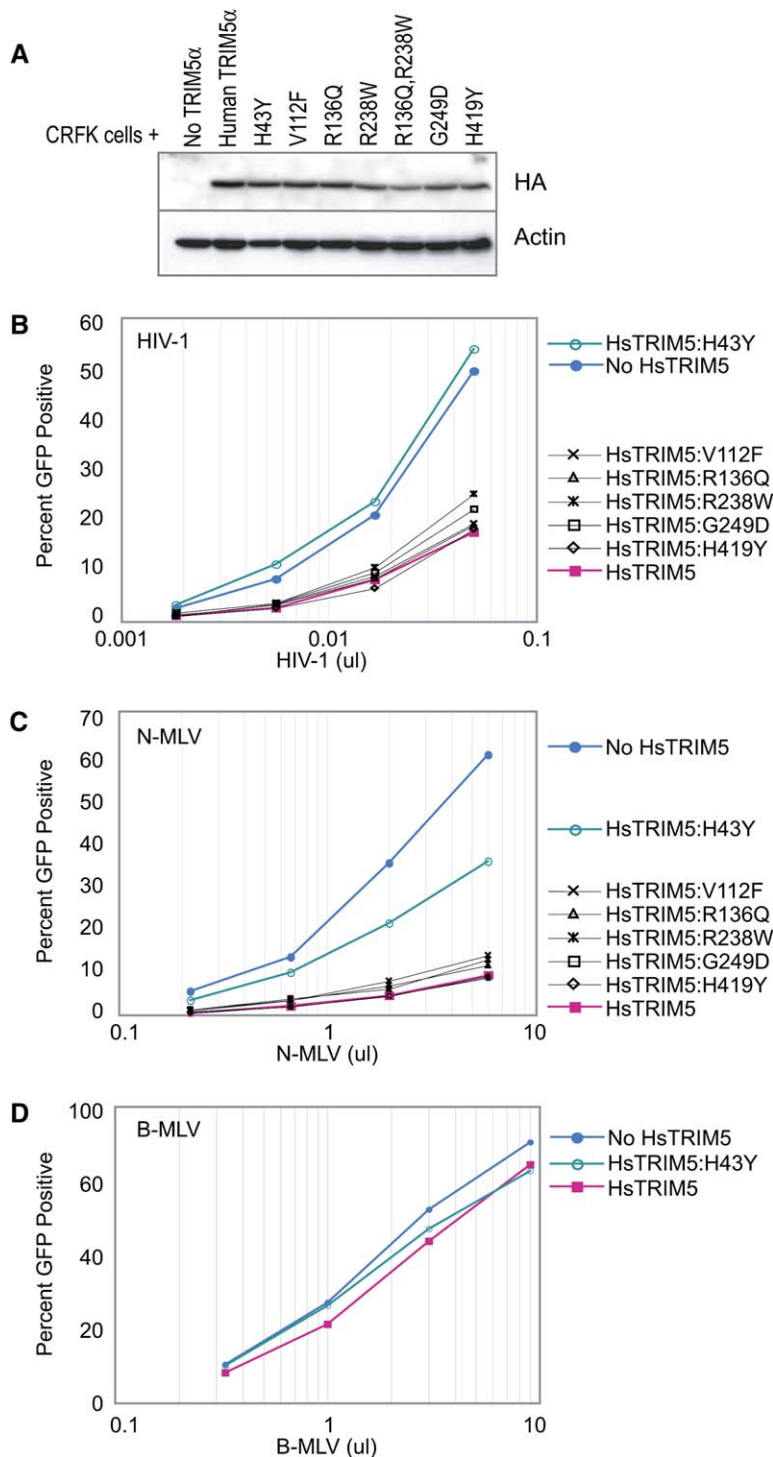


Figure 2. H43Y Negatively Impacts the Potency of Human TRIM5 α against Retroviruses
(A) A retroviral expression vector encoding human TRIM5 α was altered to contain each of the six nsSNPs. These vectors were used to stably transduce feline fibroblasts (CRFK cells), which have no intrinsic restriction to retroviruses. A Western blot shows that the HA-tagged TRIM5 α variants are expressed at equal levels. A blot with actin antibody serves as a loading control. (Based on the actin probe, H43Y may be slightly overexpressed compared to wild-type, which means that the detrimental effects of this allele are even slightly underestimated.) Because none of our human samples contained R238W in the absence of R136Q (see Figure 1A), we also tested this double mutant. Its phenotype is indistinguishable from R238W in every functional assay (data not shown). (B–D) CRFK cells (closed blue circle) or those expressing human TRIM5 α variants (all other symbols) were challenged with GFP-expressing HIV-1 (B), N-MLV (C), or B-MLV (D). H43Y (open green circle) is the only nsSNP that causes a significant change in phenotype, losing restriction to HIV-1 and N-MLV compared to wild-type human TRIM5 α (closed pink square). As expected, neither wild-type TRIM5 α nor the H43Y variant has the ability to restrict B-MLV.

human TRIM5 α has weak activity against HIV-1 when it is expressed exogenously, its endogenous expression is not a major factor in restriction of HIV-1 in human cells [17, 18].

The H43Y polymorphism changes an amino acid in the “loop 2” region of the TRIM5 α RING domain (Figure 3D), altering a position where histidine has been strictly conserved throughout primate evolution [10] (Supplemental Data). A shorter isoform of TRIM5 has been shown to have E3 ubiquitin ligase activity [19], and it has been

proposed that TRIM5 α might neutralize retroviruses through a ubiquitin-mediated pathway requiring the RING motif. This loop 2 region of a similar RING domain from human *c-cbl* has been previously implicated as the interaction interface between E2 and E3 enzymes [20]. If E2-E3 interaction were perturbed by the H43Y mutation, this would be consistent with the general (versus virus-specific) impairment of retroviral restriction that we have observed. Mutations of the invariant cysteine residues that coordinate zinc in the RING domain also

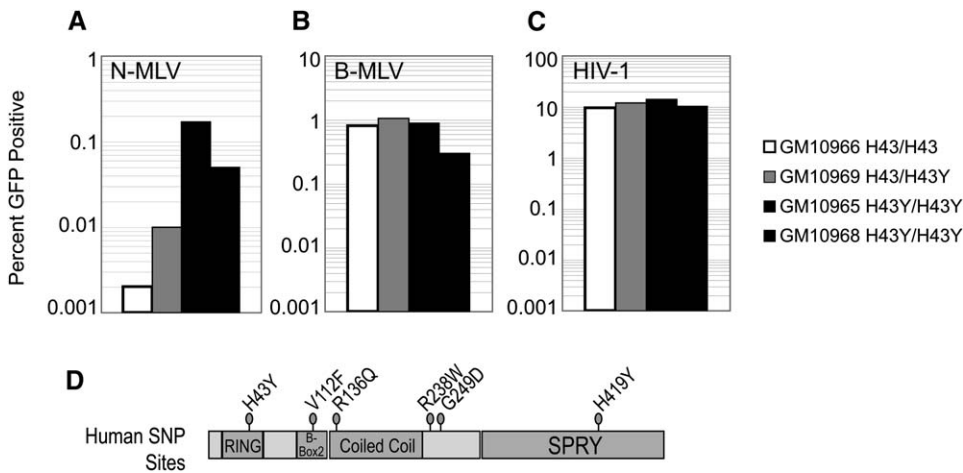


Figure 3. Restriction of Retroviruses Is Impaired in B Cells from Humans Carrying the H43Y Mutation

(A-C) B cells from four Brazilian individuals were used to test restriction of MLV and HIV in the context of human cells with different genotypes at H43. The wild-type *TRIM5α* homozygote (H43/H43, white) is very efficient at restricting N-MLV, whereas the H43Y heterozygote (H43/H43Y, gray) and homozygotes (H43Y/H43Y, black) are progressively worse at restricting N-MLV. All genotypes are essentially equal in their restriction of B-MLV and HIV-1. Two different H43Y/H43Y cell lines show consistent results. This is a representative experiment, which was performed three separate times.

(D) A schematic of the *TRIM5α* protein, with the location of the six nonsynonymous SNPs identified in our study.

significantly impair, but like H43Y, do not completely abolish restriction activity [1, 21]. We cannot formally rule out the possibility that H43Y is functionally superior at restricting a class of viruses other than that for which the wild-type variant is optimized. However, the fact that we have detected impaired function of H43Y against two distantly related retroviruses, as well as its location in the RING loop2, suggest a global loss of function.

Why has a detrimental *TRIM5α* allele persisted in the human population? We first consider H43Y prevalence in the Old World (Africa/Asia/Europe). H43Y is present on 4% of the chromosomes we sampled from the Old World (2 out of 46 chromosomes, Figure 1A). In this set, we did not observe H43Y in any of the African samples, but we did find a low incidence when we assayed additional samples representing Africans from either north or south of the Sahara (H43Y on 2 of 10 and 0 of 16 chromosomes, respectively) from the Coriell human variation panel. Consistent with H43Y origination in Africa, we estimated that the H43Y haplotype is 500,000 years old (see Supplemental Data). This compares to an estimate of 600,000 years for average, neutral human polymorphisms [22], suggesting that H43Y is not remarkable in terms of its age. We consider three models for the global persistence of H43Y in human populations: balancing selection, genetic drift, and positive selection.

First, we considered that H43Y might persist as a balanced polymorphism. Balancing selection of H43Y would carry the signature of linked, medium to high-frequency polymorphisms in the 5' region of the gene. To look for such a signature, we calculated the Tajima's D statistic, which summarizes the allele frequency distribution. For the 5' half of the gene (exons 2 to 4), which includes H43Y, we found Tajima's D to be +0.631, consistent with a trend toward medium to high-frequency SNPs, but not statistically significant for rejection of neutrality. In contrast, in the SPRY domain, Tajima's D = -0.901, in which the negative value indicates a skew toward rare alleles consistent with a recent

selective sweep in the human population, although this single polymorphism was also not sufficient for a statistically significant rejection of neutrality. However, we used the Hudson-Kreitman-Aguade (HKA) test to confirm that the polymorphism pattern of the 5' half of the gene (exons 2-4) is significantly discordant from that of the SPRY domain ($p = 0.02$). Thus, although we do not find statistical support while individually testing for either a selective sweep in the SPRY domain or balanced polymorphisms in the 5' region, these results suggest that natural selection has influenced one or both these regions during human evolution, with the most likely scenario being a selective sweep of the SPRY domain where there has been strong positive selection throughout primate evolution [10, 11]. Indeed, we find only one polymorphism, at low frequency, in the last 1.5 kb spanning exons six through eight (exon 8 encodes the SPRY domain). However, we do not find strong evidence for balancing selection acting on H43Y.

A second possibility for H43Y's persistence is a recent global relaxation of selective constraints on retroviral defense genes, resulting in this impaired allele having no contemporary fitness cost. We have previously suggested that endogenous retroviruses (ERVs) are a likely cause for the positive selection of *TRIM5α* in primates [10]. The human genome has had remarkable success in eliminating these genomic pathogens; the published human genome sequence contains thousands of ERV sequences, all of which are defective [23]. Only the HERV-K family appears to have been recently active in the human lineage [24, 25]. Furthermore, humans may have faced fewer challenges from exogenous retroviruses, as we are currently infected by only HIV and HTLV, with HIV having entered the human population only within the last century [14, 26]. Together, diminished pressure from endogenous and exogenous retroviruses in recent evolutionary history may have allowed the propagation of an impaired allele of *TRIM5* and perhaps other retroviral defense genes.

Finally, in the absence of retroviral challenge, impaired retroviral defense genes could actually have a fitness advantage. Under this model, pathogen defense genes themselves pose a selective cost to the host. For example, the *RPM1* gene from *Arabidopsis thaliana* confers resistance against *Pseudomonas* infection, yet field experiments reveal a large fitness cost imposed by *RPM1* in the absence of *Pseudomonas* [27]. In the case of *TRIM5 α* , a fitness cost in the absence of retroviruses might result from deleterious turnover of host proteins by this cytoplasmically active ubiquitin ligase, especially because *TRIM5 α* appears to be changing target specificity at a rapid evolutionary rate [10, 11]. Thus, evolutionary periods in which hosts go unchallenged by parasites can lead to the persistence or even expansion of impaired defense alleles.

Amazingly, H43Y is found at a frequency of 43% in our Central and South American samples (12/28 chromosomes), despite being detrimental to *TRIM5 α* function. This high frequency could be traced to the population bottleneck that occurred when humans migrated to the Americas with a very small effective population size, approximately 15,000 to 30,000 years ago [28]. Under this demographic model, the founder populations could have had a high frequency of H43Y and selective pressures against this allele were not stringent enough to reduce its frequency in ensuing generations.

Our study thus reveals that a significant fraction of the human population harbors an impaired variant of an important retroviral defense gene. Regardless of the selective or demographic constraints that are responsible for the current high frequency of H43Y, there may be profound future health implications of a detrimental retroviral restriction allele segregating at high frequency in the human population. Moreover, this analysis illustrates how human evolutionary history can shape our susceptibility to present-day and future viral infections.

Supplemental Data

Supplemental Data include Experimental Procedures, two figures, and one table and are available with this article online at <http://www.current-biology.com/cgi/content/full/16/1/95/DC1/>.

Acknowledgements

We thank Michael Schlador and Barbara Trask for generously sharing human DNA samples and cell lines, Richard Gardner for expert advice on RING domains, Shari Kaiser for advice on FACS analysis, and Jonathon Stoye for the gift of the N- and B-tropic MLV vectors. We thank Julie Kerns, Hua Tang, Danielle Vermaak, and three anonymous reviewers for helpful comments on the manuscript. This work was supported by National Institutes of Health grants R37 AI30927 (M.E.) and T32 CA 09657 (S.L.S.), by the Division of Nutritional Sciences and the Center for Ecogenetics and Environmental Health at the University of Washington (J.M.A.), and by startup funds from the Fred Hutchinson Cancer Research Center, a Searle Scholar Award and an Alfred P. Sloan Fellowship (H.S.M.).

Received: August 19, 2005

Revised: November 17, 2005

Accepted: November 18, 2005

Published: January 9, 2006

References

1. Stremlau, M., Owens, C.M., Perron, M.J., Kiessling, M., Autissier, P., and Sodroski, J. (2004). The cytoplasmic body component TRIM5 α restricts HIV-1 infection in Old World monkeys. *Nature* 427, 848–853.
2. Perron, M.J., Stremlau, M., Song, B., Ulm, W., Mulligan, R.C., and Sodroski, J. (2004). TRIM5 α mediates the postentry block to N-tropic murine leukemia viruses in human cells. *Proc. Natl. Acad. Sci. USA* 101, 11827–11832.
3. Hatzioannou, T., Perez-Caballero, D., Yang, A., Cowan, S., and Bieniasz, P.D. (2004). Retrovirus resistance factors Ref1 and Lv1 are species-specific variants of TRIM5 α . *Proc. Natl. Acad. Sci. USA* 101, 10774–10779.
4. Keckesova, Z., Ylisen, L.M., and Towers, G.J. (2004). The human and African green monkey TRIM5 α genes encode Ref1 and Lv1 retroviral restriction factor activities. *Proc. Natl. Acad. Sci. USA* 101, 10780–10785.
5. Yap, M.W., Nisole, S., Lynch, C., and Stoye, J.P. (2004). TRIM5 α protein restricts both HIV-1 and murine leukemia virus. *Proc. Natl. Acad. Sci. USA* 101, 10786–10791.
6. Sayah, D.M., Sokolskaja, E., Berthou, L., and Luban, J. (2004). Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV-1. *Nature* 430, 569–573.
7. Sebastian, S., and Luban, J. (2005). TRIM5 α selectively binds a restriction-sensitive retroviral capsid. *Retrovirology* 2, 40.
8. Stremlau, M., Perron, M., Welikala, S., and Sodroski, J. (2005). Species-specific variation in the B30.2(SPRY) domain of TRIM5 α determines the potency of human immunodeficiency virus restriction. *J. Virol.* 79, 3139–3145.
9. Yap, M.W., Nisole, S., and Stoye, J.P. (2005). A single amino acid change in the SPRY domain of human TRIM5 α leads to HIV-1 restriction. *Curr. Biol.* 15, 73–78.
10. Sawyer, S.L., Wu, L.I., Emerman, M., and Malik, H.S. (2005). Positive selection of primate TRIM5 α identifies a critical species-specific retroviral restriction domain. *Proc. Natl. Acad. Sci. USA* 102, 2832–2837.
11. Song, B., Gold, B., O’Huigin, C., Javanbakht, H., Li, X., Stremlau, M., Winkler, C., Dean, M., and Sodroski, J. (2005). The B30.2(SPRY) domain of the retroviral restriction factor TRIM5 α exhibits lineage-specific length and sequence variation in primates. *J. Virol.* 79, 6111–6121.
12. Perez-Caballero, D., Hatzioannou, T., Yang, A., Cowan, S., and Bieniasz, P.D. (2005). Human tripartite motif 5 α domains responsible for retrovirus restriction activity and specificity. *J. Virol.* 79, 8969–8978.
13. Nakayama, E.E., Miyoshi, H., Nagai, Y., and Shioda, T. (2005). A specific region of 37 amino acid residues in the SPRY (B30.2) domain of African green monkey TRIM5 α determines species-specific restriction of simian immunodeficiency virus SIVmac infection. *J. Virol.* 79, 8870–8877.
14. Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S., and Bhattacharya, T. (2000). Timing the ancestor of the HIV-1 pandemic strains. *Science* 288, 1789–1796.
15. Towers, G., Bock, M., Martin, S., Takeuchi, Y., Stoye, J.P., and Danos, O. (2000). A conserved mechanism of retrovirus restriction in mammals. *Proc. Natl. Acad. Sci. USA* 97, 12295–12299.
16. Andersson, M.L., Lindeskog, M., Medstrand, P., Westley, B., May, F., and Blomberg, J. (1999). Diversity of human endogenous retrovirus class II-like sequences. *J. Gen. Virol.* 80, 255–260.
17. Hatzioannou, T., Perez-Caballero, D., Cowan, S., and Bieniasz, P.D. (2005). Cyclophilin interactions with incoming human immunodeficiency virus type 1 capsids with opposing effects on infectivity in human cells. *J. Virol.* 79, 176–183.
18. Towers, G.J., Hatzioannou, T., Cowan, S., Goff, S.P., Luban, J., and Bieniasz, P.D. (2003). Cyclophilin A modulates the sensitivity of HIV-1 to host restriction factors. *Nat. Med.* 9, 1138–1143.
19. Xu, L., Yang, L., Moitra, P.K., Hashimoto, K., Rallabhandi, P., Kaul, S., Meroni, G., Jensen, J.P., Weissman, A.M., and D’Arpa, P. (2003). BTBD1 and BTBD2 colocalize to cytoplasmic bodies with the RBCC/tripartite motif protein, TRIM5 δ . *Exp. Cell Res.* 288, 84–93.
20. Zheng, N., Wang, P., Jeffrey, P.D., and Pavletich, N.P. (2000). Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* 102, 533–539.

21. Javanbakht, H., Diaz-Griffero, F., Strelau, M., Si, Z., and Sordoski, J. (2005). The contribution of RING and B-box 2 domains to retroviral restriction mediated by monkey TRIM5alpha. *J. Biol. Chem.* *280*, 26933–26940.
22. Kreitman, M., and Di Rienzo, A. (2004). Balancing claims for balancing selection. *Trends Genet.* *20*, 300–304.
23. International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature* *409*, 860–921.
24. Turner, G., Barbulescu, M., Su, M., Jensen-Seaman, M.I., Kidd, K.K., and Lenz, J. (2001). Insertional polymorphisms of full-length endogenous retroviruses in humans. *Curr. Biol.* *11*, 1531–1535.
25. Bennett, E.A., Coleman, L.E., Tsui, C., Pittard, W.S., and Devine, S.E. (2004). Natural genetic variation caused by transposable elements in humans. *Genetics* *168*, 933–951.
26. Lemey, P., Pybus, O.G., Wang, B., Saksena, N.K., Salemi, M., and Vandamme, A.M. (2003). Tracing the origin and history of the HIV-2 epidemic. *Proc. Natl. Acad. Sci. USA* *100*, 6588–6592.
27. Tian, D., Traw, M.B., Chen, J.Q., Kreitman, M., and Bergelson, J. (2003). Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature* *423*, 74–77.
28. Hey, J. (2005). On the number of New World founders: a population genetic portrait of the peopling of the Americas. *PLoS Biol.* *3*(6), e193 DOI:10.371/journal.pbio.0030193.

Accession Numbers

All sequences obtained have been deposited into Genbank under Accession numbers: [DQ298177–DQ298178](#), [DQ301444–DQ301480](#). Specifically, chimp and gorilla sequences are [DQ298177–DQ298178](#) and human sequences are [DQ301444–DQ301480](#).