

Ancient and Recent Adaptive Evolution in the Antiviral *TRIM22* Gene: Identification of a Single-Nucleotide Polymorphism That Impacts *TRIM22* Function

Jenna N. Kelly, Matthew W. Woods, Sintia Xhiku, and Stephen D. Barr*

Department of Microbiology and Immunology, Center for Human Immunology, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

Communicated by Stephen Chanock

Received 4 February 2014; accepted revised manuscript 13 May 2014.

Published online 26 May 2014 in Wiley Online Library (www.wiley.com/humanmutation). DOI: 10.1002/humu.22595

ABSTRACT: Tripartite motif protein 22 (*TRIM22*) is a novel interferon-induced protein that potently inhibits the replication of evolutionarily diverse viruses, including HIV-1. Altered *TRIM22* expression is also associated with diseases, such as multiple sclerosis, cancer, and autoimmunity. The factors that influence *TRIM22* expression and antiviral activity are largely unknown. In this study, we adopted an evolution-guided functional approach to identify potential genetic determinants of *TRIM22* function. Evolutionary analysis of *TRIM22* from mammals spanning >100 million years demonstrated that *TRIM22* evolution has been shaped by ancient and variable positive selection. We showed that positive selection is operating on multiple *TRIM22* residues that cluster in putative functional regions and that some are predicted to be functionally damaging. Interestingly, the second most prevalent *TRIM22* SNP in humans (rs1063303) is located at one of these positively selected sites. We showed that the frequency of rs1063303:G>C varies up to 10-fold between ethnicities and that in some ethnicities SNP rs1063303:G>C is being actively maintained in the population. The SNP rs1063303:G>C variant also had an inverse functional impact where it increased *TRIM22* expression and decreased the antiviral activity of *TRIM22*. Taken together, our data characterize the extensive genetic variation in *TRIM22* and identify rs1063303:G>C as a highly prevalent SNP that influences its function. Hum Mutat 00:1–10, 2014. © 2014 Wiley Periodicals, Inc.

KEY WORDS: *TRIM22*; interferon; HIV; antiviral; balancing selection; positive selection

Introduction

The tripartite motif (*TRIM*) family is a large group of proteins involved in diverse cellular processes. *TRIM* proteins are increasingly

being recognized as key regulators of the innate immune response and many *TRIMs* also have direct antiviral activity [Towers, 2007; Rajsbaum et al., 2008; Uchil et al., 2008; Carthage et al., 2009; Kajaste-Rudnitski et al., 2010; Ohmine et al., 2011]. *TRIM* proteins can be divided into two main groups based on genomic organization, domain structure, and evolutionary properties. Group 1 *TRIM* proteins (G1) are present in both invertebrate and vertebrate species and have variable C-terminal domains, whereas Group 2 *TRIM* proteins (G2) are only found in vertebrates and have a C-terminal B30.2 domain [Sardiello et al., 2008]. All *TRIM* proteins (G1 and G2) have a conserved N-terminal RBCC motif, which consists of a RING domain, one or two B-box domains, and a predicted coiled-coil region. The RING domain often has E3 ligase activity, allowing some *TRIMs* to modify other proteins (including viral proteins) with ubiquitin or ubiquitin-like molecules [Meroni and Diez-Roux, 2005; Gack et al., 2007; Duan et al., 2008].

In humans, four G2 *TRIM* genes cluster together on chromosome 11: *TRIM6*, *TRIM34*, *TRIM5*, and *TRIM22* (Online Mendelian Inheritance in Man #606559). A previous evolutionary analysis of this gene cluster showed that *TRIM6* and *TRIM34* have evolved under purifying selection in primates, whereas *TRIM5* and *TRIM22* have a dynamic evolutionary relationship that includes multiple episodes of gene expansion and loss [Sawyer et al., 2007; Tareen et al., 2009]. This study also showed that *TRIM5* and *TRIM22* have undergone positive selection in primates, with primate lineages showing positive selection in either *TRIM5* or *TRIM22*, but not in both [Sawyer et al., 2004, 2005, 2007]. In addition, *TRIM5 α* and *TRIM22* of Haplorhini primates have evolved species-specific differences in transcriptional regulation, mediated by transposable element sequences in their noncoding regions [Diehl et al., 2013]. *TRIM5 α* and *TRIM22* encode proteins that possess antiviral activity against retroviruses. The *TRIM5 α* protein inhibits HIV-1 replication in nonhuman primate cells; however, this activity is weak to absent in human cells [Stremlau et al., 2004; Song et al., 2005; Stremlau et al., 2005]. Conversely, the *TRIM22* protein inhibits HIV-1 replication in a number of human cell lines and primary monocyte-derived macrophages [Tissot and Mechti, 1995; Bouazzaoui et al., 2006; Barr et al., 2008; Kajaste-Rudnitski et al., 2011]. *TRIM22* also inhibits the replication of encephalomyocarditis virus, hepatitis B virus (HBV), and influenza A virus [Eldin et al., 2009; Gao et al., 2009; Di Pietro et al., 2013].

Genetic conflict between host and viral genomes can lead to the rapid accumulation of amino acid replacement changes (dN) relative to synonymous changes (dS). This phenomenon of positive selection (dN/dS > 1) is one hallmark of the evolutionary “battle” that occurs between host antiviral factors and their pathogen antagonists [Emerman and Malik, 2010; Duggal and Emerman, 2012]. Although positive selection is not typically observed in host genes,

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Stephen D. Barr, Western University, Dept. of Microbiology & Immunology, Dental Sciences Building Room 3006b, London, Ontario, N6A5C1, Canada. E-mail: stephen.barr@uwo.ca

Contract grant sponsors: Ontario Graduate Scholarship and Queen Elizabeth II Graduate Scholarship in Science and Technology; Canadian Institutes of Health Research (CIHR)/Canadian Association of HIV Research (CAHR); Ontario HIV Treatment Network; CIHR grants (HBF107546 and HBF134179).

multiple antiviral factors contain genetic “signatures” of positive selection [Bustamante et al., 2005; Subramanian and Kumar, 2006]. For example, well-known antiviral factors such as apolipoprotein B mRNA editing enzyme catalytic polypeptide-like (APOBEC3G), bone marrow stromal antigen 2 (BST2)/tetherin, and TRIM5 α all contain positively selected residues that play key roles in their antiviral activities [Sawyer et al., 2004, 2005; McNatt et al., 2009]. A previous evolutionary study of primate *TRIM22* sequences spanning ~33 million years of primate divergence identified a signature of positive selection for *TRIM22* in hominoids and Old World monkeys [Sawyer et al., 2007]. The positions of amino acid residues found to be under positive selection were located predominantly in the $\beta 2$ – $\beta 3$ surface loop of the B30.2 domain.

In the present study, we analyzed *TRIM22* sequences from evolutionarily diverse mammals spanning >100 million years and employed an evolution-guided functional approach to identify candidate amino acids that may influence the biological activity of human *TRIM22*. Using three-paired evolutionary models, we identified residues in several domains of *TRIM22* that are predicted to be under strong positive selection in mammals, many of which are located in or near putative functional motifs. We showed that the human *TRIM22* gene exhibits remarkable genetic diversity and contains many nonsynonymous SNPs (nsSNPs) predicted to be functionally damaging. One such nsSNP (rs1063303:G>C) is located at a site evolving under strong positive selection and is highly prevalent in the human population. Interestingly, we showed that SNP rs1063303:G>C had an inverse functional impact, where it increased *TRIM22* expression and decreased its antiviral activity. Our findings indicate that genetic variation in the *TRIM22* gene is both prevalent and ancient, and that positively selected residues in *TRIM22* can have profound influence on its antiviral activity.

Materials and Methods

Sequence Analysis

TRIM22 sequences were aligned in COBALT and a phylogenetic tree was generated using EvolView software [Papadopoulos and Agarwala, 2007; Zhang et al., 2012]. Positive selection was evaluated using the Selecton program (<http://selecton.tau.ac.il/>), as previously described [Doron-Faigenboim et al., 2005; Stern et al., 2007]. Briefly, sequences were analyzed using three-paired evolutionary models, two nested (M8a and M8; M7 and M8) and one non-nested (M8a and MEC). Models that allowed positive selection to occur (M8 and MEC) fit the data better than models that did not allow positive selection to occur (M8a and M7). The nonsynonymous (Ka) to synonymous (Ks) ratio was also calculated for each codon using Selecton. Codons with a Ka/Ks ratio >1.5 exhibited high probabilities of having evolved under positive selection.

Cells, Plasmids, and Transfections

Cells were maintained at 37°C with 5% CO₂ in standard DMEM growth medium (DMEM supplemented with 10% FBS, 100 U/ml penicillin, and 100 μ g/ml streptomycin). 293T and HeLa cell lines were obtained from the ATCC (Manassas, VA). Human osteosarcoma (HOS)-CD4⁺ cell line was obtained from NIH AIDS Reagents (Germantown, MD). The empty vector plasmid pGL3 was purchased from Promega (Madison, WI) and the pR9 plasmid encoding replication-competent HIV-1 was kindly provided by Dr. F. Bushman (University of Pennsylvania, PA). The coding region of wild-type *TRIM22* (GenBank Accession NM_006074.4) was

subcloned into p3xFLAG-CMV-10 (Sigma, St. Louis, MO) using HindIII and XbaI restriction sites to generate a plasmid encoding flag-tagged wild-type *TRIM22* (pWT-T22). The flag-tagged nsSNP rs1063303:G>C plasmid (pSNP-T22) was created using PCR mutagenesis. SNP-T22 contains C in place of the ancestral allele G at nucleotide position 725 of the *TRIM22* coding region (GenBank Accession NM_006074). The following primers were used to amplify pWT-T22 and generate PCR fragment 1: Forward WT (5' ACG TAA GCT TAT GGA TTT CTC AGT AAA GG 3') and Reverse SNP (5' GAC GAT CCC GTC AAC CTC CGC TGG AGA 3'). Similarly, PCR fragment 2 was generated using following primers: Forward SNP (5' TCT CCA GCG GAG GTT GAC GGG ATC GTC 3') and Reverse WT (5' ACG TTC TAG ATC AGG AGC TCG GTG GGC ACA CAG 3'). A 1:25 dilution of PCR fragment 1 and PCR fragment 2 was added with “Forward WT” and “Reverse WT” primers together in a PCR reaction. The amplified *TRIM22* coding region containing the nsSNP was cloned into p3xFLAG-CMV-10 (Sigma) using HindIII and XbaI to generate pSNP-T22. The entire coding region of SNP-T22 was sequenced and no other PCR-introduced variations were detected. Cells were seeded in 12-well or 6-well plates and transfected using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) with 2 or 5 μ g of plasmid DNA (pEV [empty vector control plasmid], pWT-T22, or pSNP-T22), respectively. As a control for transfection efficiency, a plasmid encoding enhanced green fluorescent protein (Clontech, Mountain View, CA) was included in transfections at a concentration one-tenth the total amount of transfected DNA. The percent GFP⁺ cells was measured using standard flow cytometry and used to control for differences in transfection efficiency. Unless otherwise stated, cotransfections of pR9 with pEV, pWT-T22, and pSNP-T22 were performed at a ratio of 10:1.

RNA Isolation and Real-Time PCR

Total RNA was extracted from cells using the R&A-BLUE Total RNA Extraction kit (Frogga Bio, Toronto, ON, Canada). To ensure that no detectable genomic or plasmid DNA was carried over in the purification, 0.1 μ g of each RNA sample was subjected to PCR (35 cycles) and real-time PCR using primers in the *TRIM22* coding region or primers specific to the *TRIM22* 3'-untranslated region. DNase-treated RNA (1 μ g) was then reverse transcribed to cDNA using the M-MLV reverse transcriptase and Oligo(dT) primers (Invitrogen). Prior to real-time PCR, cDNA samples were diluted 1:10 with water. Each PCR reaction consisted of 10 μ l of SYBR Green Master Mix, 2 μ l of the appropriate primers (1 μ l of 10 μ M forward primer [5' CAT CTG CCT GGA GCT CCT GAC 3'] and 1 μ l of 10 μ M reverse primer (5' AGA TGA TCA CTG ACT CCT TGA TCT TTG C 3')), 1 μ l of diluted cDNA, and water to a total volume of 20 μ l. Real-time PCR was run on the Rotor-Gene 6000 real-time PCR machine (Corbett Life Science, Mortlake, NSW, Australia) under the following cycling conditions: 10 min at 95°C and 40 cycles of 10 sec at 95°C, 15 sec at 60°C, and 20 sec at 72°C. The Rotor-Gene 6000 series software (version 1.7) was used to determine the C_T for each PCR reaction. All samples were amplified in triplicate with no-template controls and the mean was used for further analysis.

Western Blotting

Cells were pelleted by centrifugation (350g for 10 min), washed twice with PBS, and lysed in 1X RIPA buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1x Complete Protease Inhibitor [Roche, Basel, BS, Switzerland], 1% Triton X-100, 0.1% SDS). Virus released into the supernatant was clarified by centrifugation,

pelleted by centrifugation (21,000g for 2 hr) over a cushion of 20% sucrose and lysed with 1X RIPA buffer. Protein was separated on a 12% SDS-PAGE gel and then transferred to a FluorTransW membrane (Pall, Port Washington, NY) by semidry transfer. Following transfer, the membrane was blocked for 1 hr in LI-COR Blocking Buffer (LI-COR Biosciences, Lincoln, NE) and incubated overnight with primary antibody at 4°C (1:1,000 dilution with LI-COR Blocking Buffer). Detection was carried out using an IRDye-labeled secondary antibody (1:20,000 dilution with LI-COR Blocking Buffer for 30 min) and the LI-COR Odyssey Detection System (LI-COR Biosciences).

Confocal Immunofluorescence Microscopy

Cells were seeded in 12-well plates on 18-mm coverslips. Twenty-four hours after transfection, cells were washed with PF buffer (1x PBS + 1% FBS), fixed (1x PBS + 5% formaldehyde + 2% sucrose for 10 min), and permeabilized (1x PBS, 5% NP-40, 10% sucrose for 10 min). Cells were incubated with primary antibodies for 1 hr (1:1,000 dilution of mouse anti-Flag antibody in PF buffer), washed thoroughly with PF buffer, and then incubated with secondary antibodies for 1 hr (1:1,000 dilution of AlexaFluor 546 antimouse antibody in PF buffer). Coverslips containing the cells were mounted on glass slides with Vectashield mounting media (Vector Laboratories, Burlingame, CA) and slides were examined using a Zeiss LSM 510 confocal fluorescence microscope.

Neutrality Tests

Tajima's D and Fu's F_S neutrality tests were performed to distinguish between neutrally evolving sequences under mutation-drift equilibrium and sequences evolving under non-neutral processes such as balancing selection [Tajima, 1989a, 1989b; Fu, 1997]. These two tests are based on the principle that a recent population expansion associated with a non-neutral process will detect a shift in the allele frequency spectrum compared with a neutral Wright-Fisher model. A negative Tajima's D signifies an excess of low-frequency variants relative to expectation, indicating population size expansion (e.g., after a bottleneck or selective sweep) and/or purifying selection. A positive Tajima's D signifies low levels of both low- and high-frequency (i.e., an excess of intermediate frequency) variants, indicating a decrease in population size and/or balancing selection. A negative value of Fu's F_S is evidence for an excess number of rare alleles, as would be expected from a recent population expansion or from genetic hitchhiking. A positive value of F_S is evidence for a deficiency of alleles, as would be expected from a recent population bottleneck or from overdominant selection. The F_S statistic was considered significant at the 5% level if its P value was below 0.02 [Fu, 1997]. The analyses were performed using the software Arlequin 3.5.1.3 (Mol Ecol Res 10: 564–567). Simulated P values were generated using 10,000 simulations under a model of selective neutrality.

Statistical Analyses

Unless stated otherwise, statistical analyses were performed using GraphPad Prism version 6.0. P values of less than 0.05 were considered statistically significant.

Results

Positive Selection in Multiple TRIM22 Domains Among Mammals

To better understand the evolution of TRIM22, and to examine how selective pressures have shaped its antiviral properties, we obtained the *TRIM22* coding sequence from 29 evolutionarily diverse mammalian species (Supp. Table S1). We aligned these sequences and generated a phylogenetic tree, representing >100 million years of evolution (Fig. 1A). Our primate phylogeny was consistent with previously published studies and only orthologous nonprimate sequences were included in our analysis (Supp. Table S1) [Sawyer et al., 2007]. Using this dataset, we employed the software program Selecton to evaluate TRIM22 evolution among mammals. Selecton uses several standard evolutionary models, including two nested paired models (M8a and M8; M7 and M8) and one non-nested paired model (M8a and MEC) [Doron-Faigenboim et al., 2005; Stern et al., 2007]. We compared the nested pairs using the likelihood ratio test and found that in both cases, the model that allowed sites to evolve under positive selection (M8) fit the data significantly better than the models that did not (M8a and M7) (Table 1). The non-nested MEC model differs from the nested model in that it accounts for differences in amino acid replacement rates. A position with radical amino acid replacements will obtain a higher K_a value than a position with more moderate replacements. Akaike Information Content (AIC_c) scores are compared between the MEC and the M8a models. The lower the AIC_c score, the better the fit of the model to the data, and hence the model is considered more justified [Doron-Faigenboim et al., 2005]. Comparison of the AIC_c scores for MEC and M8a models (M8a: 25211.8; MEC: 25079.2) indicated that the data were more congruent with the model that allowed for positive selection (MEC) than the one that did not (M8a) (Table 1).

Selecton analysis identified several codons with high probabilities of having evolved under positive selection (Supp. Fig. S1). Eleven of the 28 codons predicted to be evolving under positive selection were located in the B30.2 domain (Fig. 1B; Table 2). Six of these eleven codons (K324, R327, T330, K332, S334, and C337) clustered together in one region, whereas the other five (S377, S395, G471, L488, and V489) were more dispersed. The codon numbering uses the translation initiation codon as codon 1. Interestingly, many of the positively selected codons corresponded in location and spacing to previously identified positively selected codons in the TRIM5 α protein (K324, P325, G330, R332, R335, Q337, F339, V340, K389, Q471, and G483) (Fig. 1C) [Sawyer et al., 2005]. In TRIM5 α , most positively selected codons are located in one of four variable regions (v1–v4) in the B30.2 domain. These variable regions are also found in a number of other TRIM proteins, including TRIM22 [Sawyer et al., 2005; Song et al., 2005]. The v1 region of TRIM5 α (or the “antiviral patch” region) was previously shown to be a major determinant for species-specific HIV-1 restriction in primates [Stremlau et al., 2005]. Mutations in the other variable regions (v2–v4) of TRIM5 α have also been shown to alter HIV-1, SIV, and N-MLV restriction [Yap and Stoye, 2005; Li et al., 2006; Ohkura et al., 2006; Perron et al., 2006; Kono et al., 2009]. Remarkably, all of the positively selected codons in TRIM22's B30.2 domain (except G471) were located in one of its v1–v4 regions. Six positively selected codons were found in the v1 region, two in the v2 region, and two in the v4 region. Unlike TRIM5 α , there were no positively selected codons in TRIM22's v3 region (Fig. 1D).

Outside of the B30.2 domain, we also found evidence for positively selected codons in each of TRIM22's other domains (Table 2). Many of these codons clustered together around putative functional

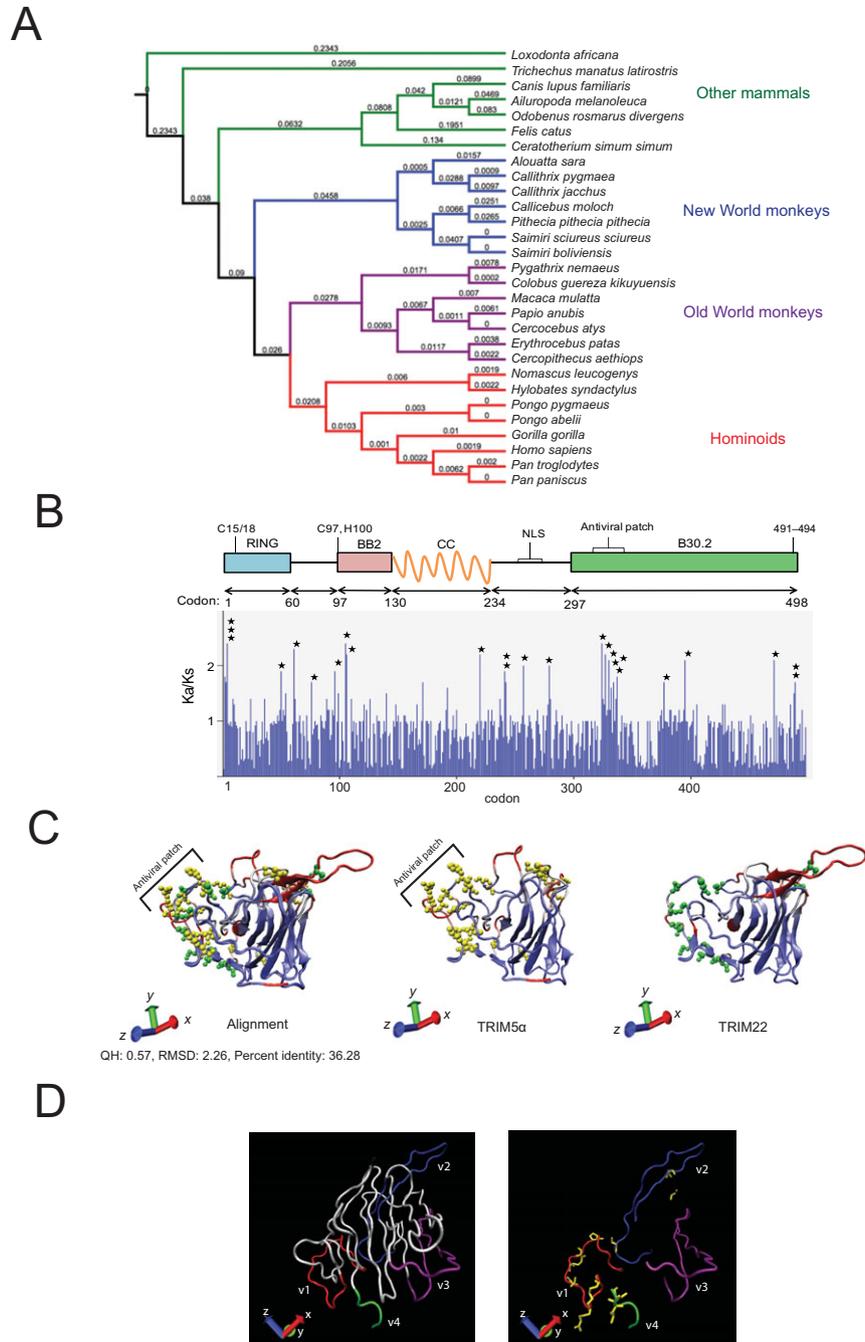


Figure 1. Mapping positively selected sites in the TRIM22 protein. **A:** Phylogenetic tree showing the evolutionary relationship among 29 mammalian species for the TRIM22 coding sequence. Tree was created using the COBALT alignment tool and EvoView software. Numbers shown represent arbitrary distances from each branch point on the tree. Hominoids are highlighted in red, Old World monkeys in purple, New World monkeys in blue, and other mammals in green. **B:** Bayesian analysis of mammalian TRIM22 coding sequences. Ka/Ks values for each codon are plotted on the y-axis. Ka/Ks ratio >1 indicates positive selection, Ka/Ks ratio <1 indicates purifying selection, and Ka/Ks ratio of 1 indicates neutral selection. Asterisks show approximate location of codons under strong positive selection. TRIM22 protein domains are shown above the graph, along with the approximate location of several functional motifs (C15/18: two cysteine residues required for E3 ligase activity; C97, H100: zinc finger motif in the B-Box 2 (BB2) domain; CC: predicted coiled-coil (CC) region; NLS: bipartite nuclear localization signal). The “antiviral patch” refers to several residues previously shown to be a major specificity determinant for TRIM5 α -mediated anti-HIV activity. **C:** Comparative model showing the putative three-dimensional structure of the B30.2 domains for both TRIM5 α and TRIM22. The blue and red colored regions represent structurally conserved regions (blue) or regions with no correspondence in structural proximities (red) after structural alignment. The Q^H value is a metric for structural homology. It is an adaptation of the Q value that measures structural conservation. $Q = 1$ implies that two structures are identical. When Q has a low score (0.1–0.3), structures are not aligned well (only a small fraction of the C_{α} atoms superimpose). Colored balls correspond to residues (van der Waals radii) predicted to evolve under strong positive selection (colored yellow for TRIM5 α and green for TRIM22). **D:** Molecular model of TRIM22's B30.2 domain. The colored regions in the image on the left show the four variable regions (v1: red; v2: blue; v3: magenta; and v4: green). The image on the right shows only the variable regions (v1–v4) plus the location of residues predicted to evolve under strong positive selection (yellow). Molecular models were created with Visual Molecular Dynamics (VMD) software (v1.9.1) support. VMD is developed with NIH support by the Theoretical and Computational Biophysics group at the Beckman Institute, University of Illinois at Urbana-Champaign.

Table 1. Models of Evolution Applied to TRIM22 Coding Sequences

Nested	Log likelihood	2 (ln ₁ –ln ₀)	Degrees of freedom	χ ²
M8a, M8	-12,601.9, -12,599.1	5.6	1	<0.001
M7, M8	-12,608.1, -12,599.1	18	2	<0.001
Non-nested	Log likelihood	AIC _c score ^{Δa}	Parameters	
M8a, MEC	-12601.9, -12534.6	25,211.8, 25,079.2	4, 5	

^{Δa}AIC_c score^Δ: $-2 \cdot \log L + 2p \cdot (N/N - p - 1)$, where L is the likelihood of the model given the data, p is the number of free parameters, and N is the sequence length. The lower the AIC_c score, the better the fit of the model to the data, and hence the model is considered more justified.

Table 2. Positive Selection in TRIM22 Protein Domains

Domain (AA ^a)	Percent PS ^b	Ka/Ks ^c	Positively selected codons
RING (1–59)	1.00	1.86	D2, F3, S4, S50, S54
SP1 (60–96)	0.60	1.97	T61, N76, V96
BB2 (97–129)	0.40	2.30	Q105, I106
CC (130–233)	0.60	1.83	A171, V192, T220
SP2 (234–296)	0.80	1.90	L241, R242, K257, R279
B30.2 (297–498)	2.21	1.90	K324, R327, T330, K332, S334, C337, S377, S395, G471, L488, V489

^aAA, amino acids included in each protein domain.

^bPercent PS, percent of positively selected sites in each domain compared to the total number of sites in the protein.

^cKa/Ks, number of nonsynonymous substitutions per nonsynonymous site (Ka) to the number of synonymous substitutions per synonymous site (Ks).

motifs. For example, three positively selected codons (V96, Q105, and I106) surrounded a zinc finger motif in the B-box 2 domain (Fig. 1B; Table 1). Zinc finger motifs are critical for proper domain folding and often contribute to protein–protein interactions [Gamsjaeger et al., 2007]. We also identified four positively selected codons (L241, R242, K257, and R279) that flanked the nuclear localization signal (NLS) in TRIM22’s Spacer 2 domain (Fig. 1B; Table 2). Along with amino acids 491–494 in the B30.2 domain, TRIM22’s NLS was previously shown to be required for its nuclear localization [Herr et al., 2009; Sivaramakrishnan et al., 2009a]. Two positively selected codons (L488 and V489) were also located directly upstream of amino acids 491–494 (Fig. 1B; Table 2).

Genetic Variation in the Human TRIM22 Gene

Since TRIM22 is an integral component of the host antiviral response, and since it has been linked to a number of diseases, we also had an interest in its evolutionary pattern among humans. To examine genetic variation in the human TRIM22 gene, we compiled a list of known nsSNPs in the TRIM22 coding region using the National Center for Biotechnology Information dbSNP database. To date, 64 nsSNPs and two indels have been identified in the exons of human TRIM22, the majority of which are located in the C-terminal B30.2 domain (Supp. Table S2). Most of these nsSNPs result in missense variations; however, several generate truncated versions of the TRIM22 protein (e.g., nonsense and frameshift variations). To identify nsSNPs that may play important roles in the activity of human TRIM22, we used a program called Sorting Intolerant From Tolerant (SIFT) to predict their functional effects [Kumar et al., 2009]. The SIFT analysis predicted that 23 nsSNPs were deleterious to TRIM22 function and 33 nsSNPs were tolerated. Paradoxically, two potentially deleterious nsSNPs (rs192306924:C>A and rs1063303:G>C)

were located at amino acid sites evolving under strong positive selection in mammals (Table 1; Supp. Table S2). Due to its high prevalence in the human population, we selected nsSNP rs1063303:G>C for further analysis and characterization.

Upon further analysis, we found that nsSNP rs1063303:G>C is caused by a guanine (G) to cytosine (C) nucleotide change, which results in an arginine (R) to threonine (T) amino acid change at position 242 in the TRIM22 protein (Fig. 2A). To determine its prevalence in the human population, we obtained frequency data from four different ethnic populations using the 1000 Genomes Database. Interestingly, we observed large differences in the frequency of nsSNP rs1063303:G>C among the various ethnic groups. For example, 31% of European individuals were homozygous for the nsSNP (C/C) genotype; however, only 3% of Asian individuals were homozygous for C/C. African and American populations had intermediate nsSNP frequencies, with 15% and 19% of individuals who were homozygous for the C/C genotype, respectively (Fig. 2B). An alignment of amino acid site 242 in mammals revealed that the ancestral allele (G) was highly conserved in hominoids and Old World monkeys, but variable in New World monkeys and other mammals (Fig. 2C). The reason for these frequency differences among humans for nsSNP rs1063303:G>C is unknown.

Tests for Selective Neutrality

A recent genome-wide scan of two ethnic populations identified TRIM22 as one of 60 “extreme” genes undergoing balancing selection in humans [Andrés et al., 2009]. To distinguish between a neutrally evolving site under mutation-drift equilibrium and a site evolving under non-neutral processes such as balancing selection, we performed Tajima’s D and Fu’s F_S tests for neutrality on the TRIM22 rs1063303 variant. These two tests are based on the principle that a recent population expansion associated with a non-neutral process will detect a shift in the allele frequency spectrum compared with a neutral Wright–Fisher model. As shown in Table 3, Tajima’s D values were positive for all populations, indicating an excess of intermediate frequency nucleotide site variants compared with the expectation under a neutral model of evolution. These deviations from neutrality were significant ($P < 0.05$) for the AFR, AMR, and EUR populations, but not for the ASN population. Similarly, the results of Fu’s F_S test also showed largely positive values for AFR, AMR, and EUR populations, indicating an excess of intermediate haplotypes over that expected under neutrality. None of the F_S coefficients were considered significant at the 5% level ($P < 0.02$); however, the F_S coefficients for AFR, AMR, and EUR populations approached significance.

Functional Analysis of the rs1063303:G>C Polymorphism

To investigate whether nsSNP rs1063303:G>C was deleterious to TRIM22 function, we transfected human cells with an empty vector control plasmid (pEV), a wild-type TRIM22 plasmid (pWT-T22), or a plasmid encoding flag-tagged TRIM22 with the rs1063303:G>C

Table 3. Results of Tajima’s D and Fu’s F_S Neutrality Tests

Population	Tajima’s D	P value	Fu’s F _S	P value
African (AFR)	2.56	0.028	2.06	0.050
Asian (ASN)	0.82	0.251	0.29	0.439
American (AMR)	2.49	0.031	1.83	0.065
European (EUR)	2.79	0.017	2.51	0.028

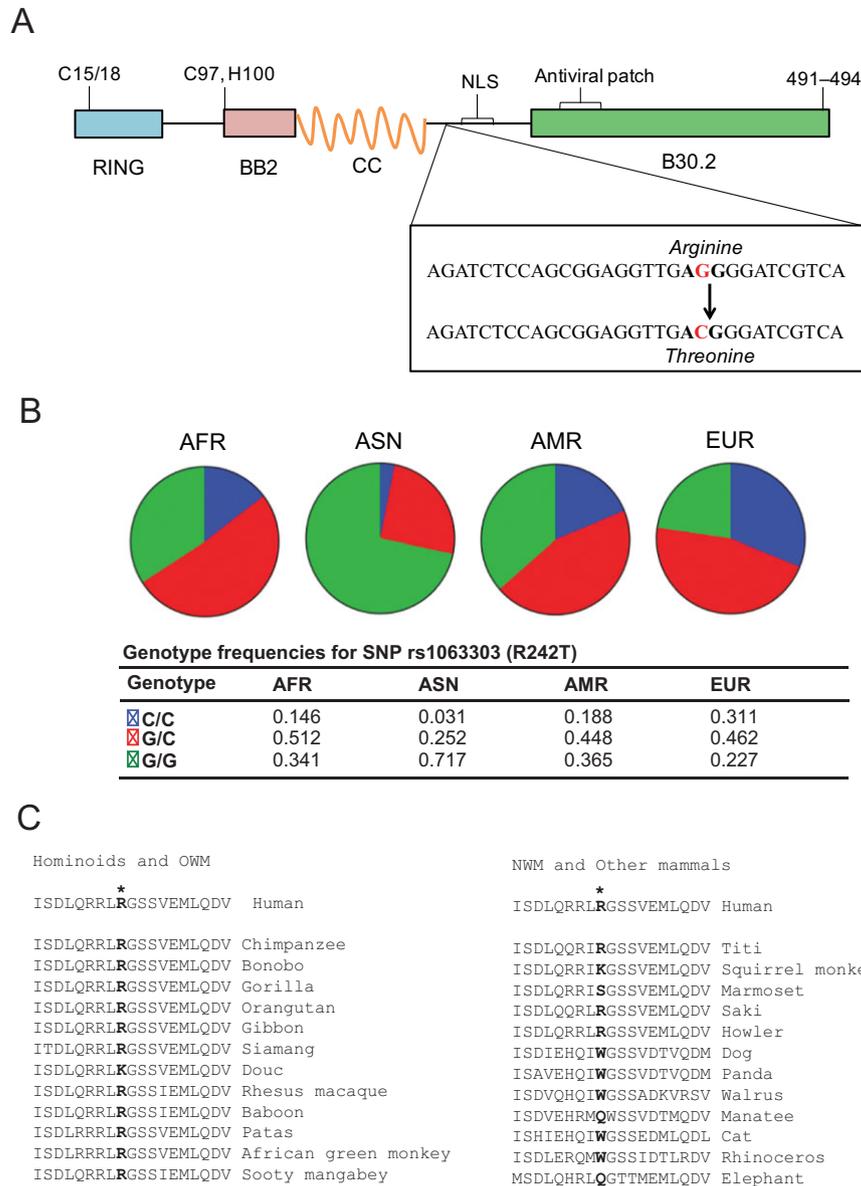


Figure 2. Genetic variation at amino acid site 242 in the TRIM22 protein. **A:** Schematic showing the approximate location of nsSNP rs1063303:G>C in the human TRIM22 protein. Inset shows the nucleotide (red) and amino acid changes for nsSNP rs1063303:G>C. C15/18: two cysteine residues required for E3 ligase activity; C97, H100: zinc finger motif in the B-Box 2 (BB2) domain; CC: predicted coiled-coil (CC) region; NLS: bipartite nuclear localization signal. The “antiviral patch” refers to several residues previously shown to be a major specificity determinant for TRIM5 α -mediated anti-HIV activity. **B:** Data were extracted from the 1000 Genomes Database. Genotype frequencies for nsSNP rs1063303:G>C in African (AFR) ($n = 246$; Yoruba in Ibadan, Nigeria; Luhya in Webuye, Kenya; Gambians in Western Divisions in The Gambia; Mende in Sierra Leone; Esan in Nigeria; Americans of African ancestry in South Western USA; African Caribbean in Barbados), American (AMR) ($n = 181$; Mexican ancestry from Los Angeles, CA; Puerto Ricans from Puerto Rico; Colombians from Medellin, Colombia; Peruvians from Lima, Peru), Asian (ASN) ($n = 286$; Han Chinese in Beijing, China; Japanese in Tokyo, Japan; Southern Han Chinese; Chinese Dai in Xishuangbanna, China; Kinh in Ho Chi Minh City, Vietnam), and European (EUR) ($n = 379$; Utah Residents (CEPH) with Northern and Western European ancestry; Toscani in Italia; Finnish in Finland; British in England and Scotland; Iberian population in Spain) populations. For each population, the percentage of wild-type homozygotes (GG), heterozygotes (GC), and nsSNP homozygotes (CC) are colored green, red, and blue, respectively. **C:** Sequence alignment of nsSNP rs1063303:G>C in hominoids and Old World monkeys (left) as well as New World monkeys and other mammals (right). Site 242 (location of nsSNP rs1063303:G>C in human TRIM22) is denoted by an asterisk and highlighted in bold.

variation (pSNP-T22). Total RNA was isolated 24 hr after transfection, reverse transcribed into cDNA and subjected to quantitative polymerase chain reaction (qPCR). Surprisingly, cells transfected with pSNP-T22 exhibited an average 40-fold increase in *TRIM22* mRNA compared with cells transfected with pWT-T22 ($P = 0.0001$, unpaired Student's *t*-test) (Fig. 3A). Whole cell lysates from similarly

transfected cells were subjected to Western blot analysis using anti-flag to analyze TRIM22 levels. Densitometric analysis of Western blots after normalization to β -actin levels and transfection efficiency revealed that TRIM22 protein levels were 10.3-fold higher in cells transfected with pSNP-T22 compared with cells transfected with pWT-T22 (Fig. 3B).

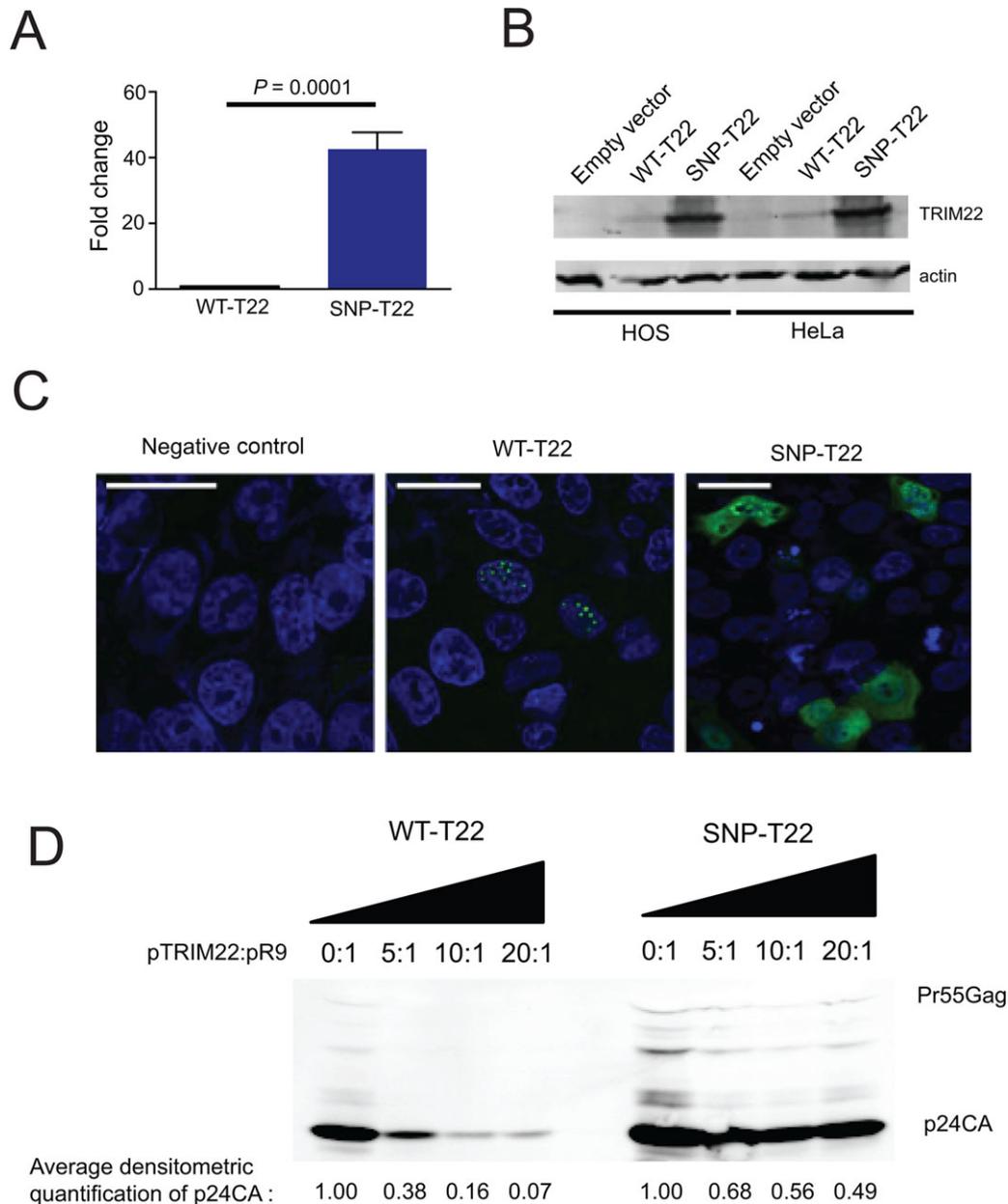


Figure 3. Identification of a nsSNP that alters TRIM22 expression and antiviral activity. **A:** Total RNA was harvested from HeLa cells transfected with empty vector pEV, plasmid expressing wild-type TRIM22 (pWT-T22) (rs1063303:G), or TRIM22 plasmid containing nsSNP rs1063303:G>C (pSNP-T22). Total RNA was extracted, reverse transcribed into cDNA and quantified using qPCR. Data are shown as the fold change relative to wild-type TRIM22 mRNA levels after normalization to β -actin levels and transfection efficiency. $P = 0.0001$ unpaired Student's *t*-test. **B:** HeLa or HOS cells were transfected with equivalent amounts of one of the following plasmids: pEV, flag-tagged pWT-T22, or flag-tagged pSNP-T22. TRIM22 protein was detected via Western blotting using an anti-flag antibody. β -Actin was used as a loading control. **C:** Confocal immunofluorescence microscopy of HeLa cells transiently transfected with pEV, pWT-T22, or pSNP-T22. HeLa cells were fixed 24 hr after transfection and TRIM22 localization was detected using an anti-flag antibody (green). Images shown represent optical slices taken through the center of cells from a series of z-stack images. DAPI staining was used to visualize the nuclei (blue). Scale bars = 10 μ m. **D:** Virus released from HeLa cells coexpressing full-length replication-competent HIV-1 (pR9) and increasing concentrations of either pWT-T22 or pSNP-T22 were resolved by SDS-PAGE and subjected to Western blot analysis using anti-p24CA. Numbers below the blots show the average densitometric quantification of p24CA after normalization to transfection efficiency. Data shown are representative of at least three independent experiments.

We also examined the effect of nsSNP rs1063303:G>C on the sub-cellular localization of TRIM22 using confocal immunofluorescence microscopy. Consistent with previous reports, WT-T22 protein localized predominantly in the nucleus and formed punctate bodies [Gao et al., 2009; Sivaramakrishnan et al., 2009b; Sivaramakrishnan et al., 2009a] (Fig. 3C). In contrast, SNP-T22 protein localized diffusely in both the cytoplasm and the nucleus. Previous studies have

demonstrated that TRIM22 can inhibit HIV-1 particle production in human cells [Tissot and Mechetti, 1995; Bouazzaoui et al., 2006; Barr et al., 2008; Kajaste-Rudnitski et al., 2011]. To assess the impact of nsSNP rs1063303:G>C on the antiviral activity of TRIM22, HOS cells or HeLa cells were cotransfected with plasmids encoding replication-competent proviral HIV-1 (pR9) and either WT-T22 or SNP-T22. HIV-1 Gag is expressed as a Pr55Gag polyprotein, which

consists of an amino-terminal Matrix protein (17 kDa), followed by Capsid protein (“p24CA”) (24 kDa), p2 protein (2 kDa), nucleocapsid protein (7 kDa), p1 protein (1 kDa), and p6 protein (6 kDa) at the carboxyl-terminus. Concomitant with HIV release, Gag is proteolytically cleaved by HIV-1 protease to yield each of these individual proteins. Intracellular Gag protein level was measured by quantitative Western blotting using anti-p24CA. Cells expressing increasing concentrations of WT-T22 exhibited decreasing levels of intracellular HIV-1 Gag protein expression. In contrast, cells expressing increasing concentrations of SNP-T22 did not exhibit a substantial reduction in intracellular Gag protein production. Taken together, these data showed that the SNP rs1063303:G>C variant had an inverse functional impact where it increased TRIM22 expression and decreased the antiviral activity of TRIM22.

Discussion

Recent studies on host restriction factors have used an evolutionary approach to identify amino acid residues that are required for their antiviral activity. These studies exploit the evolutionary “arms race” that occurs between host restriction factors and viruses as they compete to gain an evolutionary advantage over each other. As a result of this evolutionary “battle,” many host restriction factors contain genetic signatures of positive selection, particularly at amino acid sites that interact with viral antagonists [Sawyer et al., 2004; Sawyer et al., 2005; McNatt et al., 2009]. Here, we used a similar evolutionary approach to analyze *TRIM22* sequences spanning >100 million years of evolution. In addition to the positively selected residues previously identified in the coiled-coil and B30.2 domains [Sawyer et al., 2007], we identified positively selected residues in several other domains of TRIM22. We also showed that the human *TRIM22* gene contains multiple nsSNPs with the potential to alter TRIM22 function and identified a highly prevalent nsSNP in *TRIM22* that alters its expression and antiviral activity against HIV-1 replication.

Positively selected residues are often located near protein-binding sites and tend to be solvent exposed. In addition, these residues typically occur in clusters with other positively selected residues [Lin et al., 2007]. Although the tertiary structure of TRIM22 has not yet been resolved, its primary amino acid sequence contains a number of putative protein-binding motifs. For example, RING domains with E3 ligase activity-bind E2-conjugating enzymes via their zinc finger motifs to mediate ubiquitin transfer to a substrate protein [Gamsjaeger et al., 2007]. TRIM22 contains a zinc finger motif in its RING domain that facilitates E3 ligase activity in combination with the E2 enzyme UbcH5b [Duan et al., 2008]. TRIM22 also contains a second zinc finger motif in its B-box 2 domain, a bipartite NLS in its Spacer 2 domain, and a short sequence in its B30.2 domain that is required for its nuclear localization (amino acids 491–494) [Herr et al., 2009; Sivaramakrishnan et al., 2009a]. Remarkably, 13 of the 28 positively selected residues that we identified in TRIM22 are located within 15 amino acids of one of these motifs (Fig. 1B). Most are not located directly within the motifs, but instead cluster around them in groups of three or four. Although in TRIM22, none of these motifs have been explicitly shown to interact with other proteins, this type of evolutionary pattern suggests that they may indeed function as protein-binding sites.

We identified a number of positively selected residues in the B30.2 domain of TRIM22 that correspond in location and spacing to positively selected residues in TRIM5 α . In addition, similar to TRIM5 α , we showed that the majority of these residues are located in one of four variable regions (v1–v4). In rhesus monkey TRIM5 α , v1–v4

form flexible loops that map to the structurally divergent face of the protein [Biris et al., 2012]. Multiple residues within TRIM5 α v1–v4 regions are critical for virus restriction, including several positively selected sites. For example, residues 324 and 332 in the v1 region and residues 385 and 389 in the v2 region are required for TRIM5 α -mediated inhibition of HIV-1 and/or SIV in hominoids [Stremlau et al., 2005; Ohkura et al., 2006; Kono et al., 2009]. Moreover, in the v3 region, residues 409 and 410 are required for N-MLV restriction by human TRIM5 α [Perron et al., 2006]. Unlike TRIM5 α , there are no residues evolving under positive selection in the v3 region of TRIM22. Thus, it is possible that this region of TRIM22 is not subject to strong evolutionary pressures (e.g., from viral antagonists). Consistent with this, TRIM22 does not inhibit N-MLV replication in human cells [Barr et al., 2008]. Similar to TRIM5 α , TRIM22 has been shown to interact with the HIV-1 Gag and/or capsid protein [Sebastian and Luban, 2005; Barr et al., 2008; Pertel et al., 2011]. Because the v1–v4 regions of rhesus monkey TRIM5 α form an extensive HIV-1 capsid-binding interface, it is possible that these regions are also important for TRIM22 and HIV-1 Gag/capsid binding. It will be interesting to learn whether the v1, v2, and v4 regions of TRIM22 are targeted by an HIV-1 antagonist, thereby placing evolutionary pressure on TRIM22 to, for example, maintain its interaction with the HIV-1 Gag protein.

Genes that evolve under positive selection during interspecies evolution also tend to be highly polymorphic in humans [Liu et al., 2008]. We demonstrate here that the human *TRIM22* gene contains multiple nsSNPs and that many of these nsSNPs are predicted to be deleterious to protein function. We characterized one potentially deleterious nsSNP (rs1063303:G>C) that is highly prevalent in the human population and show that it is located at an amino acid site in TRIM22 that has undergone strong positive selection in mammals. Of interest, the frequency of nsSNP rs1063303:G>C varies considerably among different ethnic populations. Ethnic differences in nsSNP frequencies have been reported for many genes, including several nsSNPs in toll-like receptor genes and a number nsSNPs associated with autoimmune disease [Mori et al., 2005; Cheng et al., 2007]. It is possible that the differences in nsSNP rs1063303:G>C frequency are due to differential prevalence of certain diseases in different geographic locations. For example, one nsSNP in the β -globin gene is highly prevalent in regions that are endemic for malaria, but not in nonendemic regions. This nsSNP produces an abnormal version of the β -globin protein called hemoglobin S (HbS), which causes sickle-cell disease in homozygotes (HbSS), but affords protection from malaria in heterozygotes (HbAS) [Rees et al., 2010]. As such, the “deleterious” HbS allele is maintained at higher than expected frequencies in certain populations. This type of evolution, whereby polymorphism is maintained in a population because it confers a selective advantage, is referred to as balancing selection [Bustamante et al., 2005].

Several host restriction factors have been shown to undergo balancing selection in primates, including human TRIM5 α and primate OAS1 [Newman et al., 2006; Compton et al., 2012; Ferguson et al., 2012]. Moreover, a recent genome-wide scan of two ethnic populations identified TRIM22 as one of 60 “extreme” genes undergoing balancing selection in humans [Andrés et al., 2009]. Given its high frequency in certain human populations and history of positive selection in mammals, it is possible that nsSNP rs1063303 confers some selective advantage in heterozygotes and is maintained by balancing selection in these populations. For example, the nsSNP allele may be more advantageous against a virus other than HIV-1 that is particularly prevalent in certain geographic locations. The overall positive values resulting from the Tajima's D and Fu's F_s tests suggest that there is an excess of intermediate frequency of the *TRIM22*

rs1063303 alleles, which can imply a decrease in population size and/or balancing selection. Together, our analyses reveal a complex and multifaceted scenario for the evolution of nsSNP rs1063303 and the *TRIM22* gene.

Other retroviral restriction factors, such as APOBEC3G and TRIM5 α , contain nsSNPs that alter their antiviral activity against HIV-1. For example, the H186R variant in APOBEC3G is strongly associated with CD4⁺ T-cell decline and accelerated disease progression in African Americans. Interestingly, this association is not present in Caucasian individuals or Europeans [An et al., 2004; Do et al., 2005; Reddy et al., 2010]. A number of TRIM5 α nsSNPs, including H43Y, R136Q, and G249D, also correlate with notable differences in HIV-1 acquisition and disease progression [Sawyer et al., 2006; Sobieszczyk et al., 2011]. We showed here that a TRIM22 clone containing the ancestral allele rs1063303:G potentially inhibited particle production of full-length replication-competent HIV-1, whereas a TRIM22 clone containing the derived allele rs1063303:C drastically increased the amount of *TRIM22* mRNA and protein in human cells, altered its subcellular localization, and failed to inhibit HIV-1 particle production. This finding contrasts a previous finding by Ghezzi et al. (2013) who showed that a *TRIM22* clone containing the derived allele rs1063303:C inhibited expression of luciferase from a reporter construct containing luciferase under control from the HIV-1 long terminal repeat (LTR) promoter. A likely reason for this difference could be attributed to differences between the two systems where we used full-length HIV-1 and they used only the HIV-1 LTR, implying that other HIV-1 proteins (e.g., antagonists) may affect the antiviral activity of TRIM22. In addition, Ghezzi et al. 2013 showed that SNP rs1063303 alone was not associated with disease progression; however, a TRIM22 haplotype involving specific SNP alleles of rs1063303 and rs7935564 was found more frequently in advanced progressors than in long-term nonprogressors. Together, these findings highlight the importance of function-altering *TRIM22* SNPs and haplotypes, warranting further investigation into their clinical significance.

It remains unclear how TRIM22 rs1063303:G>C increases *TRIM22* mRNA and protein levels. nsSNPs in the gene coding region are typically expected to alter protein function, expression, conformation, or stability. It is possible that the amino acid change associated with rs1063303 may affect TRIM22 protein stability or its ability to undergo self-ubiquitination and proteasomal degradation. TRIM22 has been shown to inhibit expression from the promoters of HIV-1 and HBV [Tissot and Mechti, 1995; Gao et al., 2009; Kajaste-Rudnitski et al., 2011]. The different TRIM22 SNP rs1063303 variants may exhibit reduced or enhanced TRIM22 activity, thereby affecting gene expression from TRIM22 target promoters such as the HIV-1 LTR and/or the promoter controlling expression of TRIM22 itself. Other explanations for the observed difference in TRIM22 RNA and protein levels include, among others, structural changes in the RNA transcript that can in turn influence splicing, stability, or translational regulation. Further studies are required to determine precisely how TRIM22 rs1063303:G>C affects TRIM22 mRNA and protein levels.

Interestingly, *TRIM22* mRNA expression levels have previously been shown to influence HIV infection in vivo. In the Centre for the AIDS Programme of Research in South Africa (CAPRISA) study cohort, expression of *TRIM22* mRNA was positively correlated with CD4⁺ T-cell count and negatively correlated with viral load [Singh et al., 2011]. Conversely, a study in the Swiss HIV study cohort found that expression of *TRIM22* mRNA was positively correlated with HIV-1 RNA levels at the viral set point [Rotger et al., 2010]. Given the differences in nsSNP rs1063303 frequency among different ethnic populations, it is interesting to note that the CAPRISA

cohort comprised entirely black African females, whereas the Swiss cohort comprised entirely Caucasian individuals. It is tempting to speculate that the high prevalence of nsSNP rs1063303:G>C in Caucasians is at least partially responsible for the increased viral load in these individuals. Further studies that include populations from different geographic regions are needed to determine whether and how nsSNP rs1063303 influences HIV-1 infection in these different groups and whether other SNPs are involved.

Previous studies on host restriction factors have shown that they are rapidly evolving genes that must remain genetically “flexible” to adapt to changing pathogenic landscapes. Although their flexibility is integral to host survival, it can also be detrimental to protein function. Our study characterizes the effects of nsSNP rs1063303:G>C on TRIM22 protein function and provides insight into the ancient and extensive genetic variation within the *TRIM22* gene.

Acknowledgments

We thank Dr. John Koval for helpful discussions about neutrality tests.

Disclosure statement: The authors have no conflict of interest to declare.

References

- An P, Bleiber G, Duggal P, Nelson G, May M, Mangeat B, Alobwede I, Trono D, Vlahov D, Donfield S, Goedert JJ, Phair J, et al. 2004. APOBEC3G genetic variants and their influence on the progression to AIDS. *J Virol* 78:11070–11076.
- Andrés AM, Hubisz MJ, Indap A, Torgerson DG, Degenhardt JD, Boyko AR, Gutenkunst RN, White TJ, Green ED, Bustamante CD, Clark AG, Nielsen R. 2009. Targets of balancing selection in the human genome. *Mol Biol Evol* 26:2755–2764.
- Barr SD, Smiley JR, Bushman FD. 2008. The interferon response inhibits HIV particle production by induction of TRIM22. *PLoS Pathog* 4:e1000007.
- Biris N, Yang Y, Taylor AB, Tomashevski A, Guo M, Hart PJ. 2012. Structure of the rhesus monkey TRIM5 α PRYSPRY domain, the HIV capsid recognition module. *Proc Natl Acad Sci USA* 109:13278–13283.
- Bouazzaoui A, Kreutz M, Eisert V, Dinauer N, Heinzelmann A, Hallenberger S, Straley J, Walker R, Rübsamen-Waigmann H, Andreesen R, von Briesen H. 2006. Stimulated trans-acting factor of 50 kDa (Staf50) inhibits HIV-1 replication in human monocyte-derived macrophages. *Virology* 356:79–94.
- Bustamante CD, Fledel-Alon A, Williamson S, Nielsen R, Hubisz MT, Glanowski S, Tanenbaum DM, White TJ, Sninsky JJ, Hernandez RD, Civello D, Adams MD, et al. 2005. Natural selection on protein-coding genes in the human genome. *Nature* 437:1153–1157.
- Carthage L, Bergamaschi A, Luna JM, David A, Uchil PD, Margottin-Goguet F, Mothes W, Hazan U, Transy C, Pancino G, Nisole S. 2009. Human TRIM gene expression in response to interferons. *PLoS One* 4:e4894.
- Cheng P-L, Eng H-L, Chou M-H, You H-L, Lin T-M. 2007. Genetic polymorphisms of viral infection-associated Toll-like receptors in Chinese population. *Transl Res* 150:311–318.
- Compton AA, Hirsch VM, Emerman M. 2012. The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict. *Cell Host Microbe* 11:91–98.
- Diehl W, Johnson W, Hunter E. 2013. Elevated rate of fixation of endogenous retroviral elements in Haplorhini TRIM5 and TRIM22 genomic sequences: impact on transcriptional regulation. *PLoS One* 8:e58532.
- Do H, Vasilescu A, Diop G, Hirtzig T, Heath SC, Coulonges C, Rappaport J, Therwath A, Lathrop M, Matsuda F, Zagury J-F. 2005. Exhaustive genotyping of the CEM15 (APOBEC3G) gene and absence of association with AIDS progression in a French cohort. *J Infect Dis* 191:159–163.
- Doron-Faigenboim A, Stern A, Mayrose I, Bacharach E, Pupko T. 2005. Selecton: a server for detecting evolutionary forces at a single amino-acid site. *Bioinformatics* 21:2101–2103.
- Duan Z, Gao B, Xu W, Xiong S. 2008. Identification of TRIM22 as a RING finger E3 ubiquitin ligase. *Biochem Biophys Res Commun* 374:502–506.
- Duggal NK, Emerman M. 2012. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat Rev Immunol* 12:687–695.
- Eldin P, Papon L, Oteiza A, Brocchi E, Lawson TG, Mechti N. 2009. TRIM22 E3 ubiquitin ligase activity is required to mediate antiviral activity against encephalomyocarditis virus. *J Gen Virol* 90:536–545.
- Emerman M, Malik HS. 2010. Paleovirology—modern consequences of ancient viruses. *PLoS Biol* 8:e1000301.

- Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925.
- Ferguson W, Dvora S, Fikes RW, Stone AC, Boissinot S. 2012. Long-term balancing selection at the antiviral gene OAS1 in Central African chimpanzees. *Mol Biol Evol* 29:1093–1103.
- Gack MU, Shin YC, Joo C-H, Urano T, Liang C, Sun L, Takeuchi O, Akira S, Chen Z, Inoue S, Jung JU. 2007. TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* 446:916–920.
- Gamsjaeger R, Liew CK, Loughlin FE, Crossley M, Mackay JP. 2007. Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends Biochem Sci* 32:63–70.
- Gao B, Duan Z, Xu W, Xiong S. 2009. Tripartite motif-containing 22 inhibits the activity of hepatitis B virus core promoter, which is dependent on nuclear-located RING domain. *Hepatology* 50:424–433.
- Ghezzi S, Galli L, Kajaste-Rudnitski A, Turrini F, Marelli S, Toniolo D, Casoli C, Riva A, Poli G, Castagna A, Vicenzi E. 2013. Identification of TRIM22 single nucleotide polymorphisms associated with loss of inhibition of HIV-1 transcription and advanced HIV-1 disease. *AIDS* 27:2335–2344.
- Herr A-M, Dressel R, Walter L. 2009. Different subcellular localisations of TRIM22 suggest species-specific function. *Immunogenetics* 61:271–280.
- Kajaste-Rudnitski A, Marelli SS, Pultrone C, Pertel T, Uchil PD, Mechti N, Mothes W, Poli G, Luban J, Vicenzi E. 2011. TRIM22 inhibits HIV-1 transcription independently of its E3 ubiquitin ligase activity, Tat, and NF-kappaB-responsive long terminal repeat elements. *J Virol* 85:5183–5196.
- Kajaste-Rudnitski A, Pultrone C, Marzetta F, Ghezzi S, Coradin T, Vicenzi E. 2010. Restriction factors of retroviral replication: the example of Tripartite Motif (TRIM) protein 5 alpha and 22. *Amino Acids* 39:1–9.
- Reddy K, Winkler CA, Werner L, Mlisana K, Abdool Karim SS, Ndung'u T; CAPRISA Acute Infection Study Team. 2010. APOBEC3G expression is dysregulated in primary HIV-1 infection and a polymorphic variant influences CD4+ T-cell counts and plasma viral load. *AIDS* 24:195–204.
- Kono K, Bozek K, Domingues FS, Shioda T, Nakayama EE. 2009. Impact of a single amino acid in the variable region 2 of the Old World monkey TRIM5alpha SPRY (B30.2) domain on anti-human immunodeficiency virus type 2 activity. *Virology* 388:160–168.
- Kumar P, Henikoff S, Ng PC. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4:1073–1081.
- Li Y, Li X, Stremlau M, Lee M, Sodroski J. 2006. Removal of arginine 332 allows human TRIM5alpha to bind human immunodeficiency virus capsids and to restrict infection. *J Virol* 80:6738–6744.
- Lin Y-S, Hsu W-L, Hwang J-K, Li W-H. 2007. Proportion of solvent-exposed amino acids in a protein and rate of protein evolution. *Mol Biol Evol* 24:1005–1011.
- Liu J, Zhang Y, Lei X, Zhang Z. 2008. Natural selection of protein structural and functional properties: a single nucleotide polymorphism perspective. *Genome Biol* 9:R69.
- McNatt MW, Zang T, Hatziioannou T, Bartlett M, Fofana I Ben, Johnson WE, Neil SJD, Bieniasz PD. 2009. Species-specific activity of HIV-1 Vpu and positive selection of tetherin transmembrane domain variants. *PLoS Pathog* 5:e1000300.
- Meroni G, Diez-Roux G. 2005. TRIM/RBCC, a novel class of “single protein RING finger” E3 ubiquitin ligases. *Bioessays* 27:1147–1157.
- Mori M, Yamada R, Kobayashi K, Kawaida R, Yamamoto K. 2005. Ethnic differences in allele frequency of autoimmune-disease-associated SNPs. *J Hum Genet* 50:264–266.
- Newman RM, Hall L, Connole M, Chen G-L, Sato S, Yuste E, Diehl W, Hunter E, Kaur A, Miller GM, Johnson WE. 2006. Balancing selection and the evolution of functional polymorphism in Old World monkey TRIM5alpha. *Proc Natl Acad Sci USA* 103:19134–19139.
- Ohkura S, Yap MW, Sheldon T, Stoye JP. 2006. All three variable regions of the TRIM5alpha B30.2 domain can contribute to the specificity of retrovirus restriction. *J Virol* 80:8554–8565.
- Ohmine S, Sakuma R, Sakuma T, Thatava T, Takeuchi H, Ikeda Y. 2011. The antiviral spectra of TRIM5alpha orthologues and human TRIM family proteins against lentiviral production. *PLoS One* 6:e16121.
- Papadopoulos JS, Agarwala R. 2007. COBAL: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* 23:1073–1079.
- Perron MJ, Stremlau M, Sodroski J. 2006. Two surface-exposed elements of the B30.2/SPRY domain as potency determinants of N-tropic murine leukemia virus restriction by human TRIM5alpha. *J Virol* 80:5631–5636.
- Pertel T, Hausmann S, Morger D, Züger S, Guerra J, Lascano J, Reinhard C, Santoni F, Uchil PD, Bisiaux A, Albert M, Strambio-de-castillia C, et al. 2011. TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature* 472:361–365.
- Pietro A Di, Kajaste-Rudnitski A, Oteiza A, Nicora L, Towers GJ, Mechti N, Vicenzi E. 2013. TRIM22 inhibits influenza A virus infection by targeting the viral nucleoprotein for degradation. *J Virol* 87:4523–4533.
- Rajsbaum R, Stoye JP, O’Garra A. 2008. Type I interferon-dependent and -independent expression of tripartite motif proteins in immune cells. *Eur J Immunol* 38:619–630.
- Rees DC, Williams TN, Gladwin MT. 2010. Sickle-cell disease. *Lancet* 376:2018–2031.
- Rotger M, Dang KK, Fellay J, Heinzen EL, Feng S, Descombes P, Shianna K V, Ge D, Günthard HF, Goldstein DB, Telenti A. 2010. Genome-wide mRNA expression correlates of viral control in CD4+ T-cells from HIV-1-infected individuals. *PLoS Pathog* 6:e1000781.
- Sardiello M, Cairo S, Fontanella B, Ballabio A, Meroni G. 2008. Genomic analysis of the TRIM family reveals two groups of genes with distinct evolutionary properties. *BMC Evol Biol* 8:225.
- Sawyer SL, Emerman M, Malik HS. 2004. Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. *PLoS Biol* 2:E275.
- Sawyer SL, Wu LI, Emerman M, Malik HS. 2005. Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. *Proc Natl Acad Sci USA* 102:2832–2837.
- Sawyer SL, Wu LI, Akey JM, Emerman M, Malik HS. 2006. High-frequency persistence of an impaired allele of the retroviral defense gene TRIM5alpha in humans. *Curr Biol* 16:95–100.
- Sawyer SL, Emerman M, Malik HS. 2007. Discordant evolution of the adjacent antiviral genes TRIM22 and TRIM5 in mammals. *PLoS Pathog* 3:e197.
- Sebastian S, Luban J. 2005. TRIM5alpha selectively binds a restriction-sensitive retroviral capsid. *Retrovirology* 2:40.
- Singh R, Gaiha G, Werner L, McKim K, Mlisana K, Luban J, Walker BD, Karim SSA, Brass AL, Ndung’u T. 2011. Association of TRIM22 with the type 1 interferon response and viral control during primary HIV-1 infection. *J Virol* 85:208–216.
- Sivaramakrishnan G, Sun Y, Tan SK, Lin VCL. 2009a. Dynamic localization of tripartite motif-containing 22 in nuclear and nucleolar bodies. *Exp Cell Res* 315:1521–1532.
- Sivaramakrishnan G, Sun Y, Rajmohan R, Lin VCL. 2009b. B30.2/SPRY domain in tripartite motif-containing 22 is essential for the formation of distinct nuclear bodies. *FEBS Lett* 583:2093–2099.
- Sobieszczyk ME, Lingappa JR, McElrath MJ. 2011. Host genetic polymorphisms associated with innate immune factors and HIV-1. *Curr Opin HIV AIDS* 6:427–434.
- Song B, Gold B, Colm O, Javanbakht H, Li X, Stremlau M, Winkler C, Dean M, Sodroski J. 2005. The B30.2 (SPRY) domain of the retroviral restriction factor TRIM5alpha exhibits lineage-specific length and sequence variation in primates. *J Virol* 79:6111–6121.
- Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T. 2007. Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res* 35:W506–W511.
- Stremlau M, Owens CM, Perron MJ. 2004. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature* 427:848–853.
- Stremlau M, Perron M, Welikala S, Sodroski J. 2005. Species-specific variation in the B30.2 (SPRY) domain of TRIM5alpha determines the potency of human immunodeficiency virus restriction. *J Virol* 79:3139–3145.
- Subramanian S, Kumar S. 2006. Higher intensity of purifying selection on >90% of the human genes revealed by the intrinsic replacement mutation rates. *Mol Biol Evol* 23:2283–2287.
- Tajima F. 1989a. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Tajima F. 1989b. The effect of change in population size on DNA polymorphism. *Genetics* 123:597–601.
- Tareen S, Sawyer S, Malik H, Emerman M. 2009. An expanded clade of rodent Trim5 genes. *Virology* 385:473–483.
- Tissot C, Mechti N. 1995. Molecular cloning of a new interferon-induced factor that represses human immunodeficiency virus type 1 long terminal repeat expression. *J Biol Chem* 270:14891–14898.
- Towers GJ. 2007. The control of viral infection by tripartite motif proteins and cyclophilin A. *Retrovirology* 4:40.
- Uchil PD, Quinlan BD, Chan W-T, Luna JM, Mothes W. 2008. TRIM E3 ligases interfere with early and late stages of the retroviral life cycle. *PLoS Pathog* 4:e16.
- Yap MW, Stoye JP. 2005. A Single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction. *Curr Biol* 15:73–78.
- Zhang H, Gao S, Lercher MJ, Hu S, Chen W-H. 2012. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res* 40:W569–W572.