Chance favors a prepared genome

Jonathan P. Stoye* and Melvyn W. Yap

Division of Virology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom

During evolution, all vertebrates have been exposed to multiple waves of cross-species infection by retroviruses. A number of defensive strategies have developed, among them the spontaneous expression, throughout the body, of restriction factors, antiviral proteins that will bind to the capsid (CA) protein of incoming virions and interfere with virus replication (1). One example of such a protein is TRIM5α, first described in 2004, that provides species-specific retrovirus restriction and can prevent HIV-1 infection (2). TRIM5α can block retrovirus replication either before or after reverse transcription seemingly dependent on proteasome engagement (3). It comprises an N-terminal RBCC domain made up of RING (R), B-box (B), and coiled-coil (CC) modules and a C-terminal B30.2 domain (4). The B30.2 domain mediates virus binding and determines which viruses a given TRIM5α will restrict, whereas the RBCC domain is important for TRIM5α multimerization and appears responsible for communicating with the proteasome. Shortly after the initial description of Trim5α, a novel restriction protein was reported in *Aotus trivirgatus* (owl monkey). Insertion of a cyclophilin A (CypA) pseudogene between exons 7 and 8 of the TRIM5 gene resulted in formation of a fusion protein where the CypA has replaced the B30.2 domain (5, 6). Because CypA can bind CA from various lentiviruses (7), this fusion protein can act as a restriction factor for a number of viruses including HIV-1. Remarkably, it now appears that the owl monkey fusion protein is not unique. Three articles in this issue of PNAS, by Wilson *et al.* (8), Virgen *et al*. (9), and Brennan *et al.* (10), and two articles published elsewhere (11, 12) provide a description of the second, independent case of the generation of a Trim5CypA fusion protein.

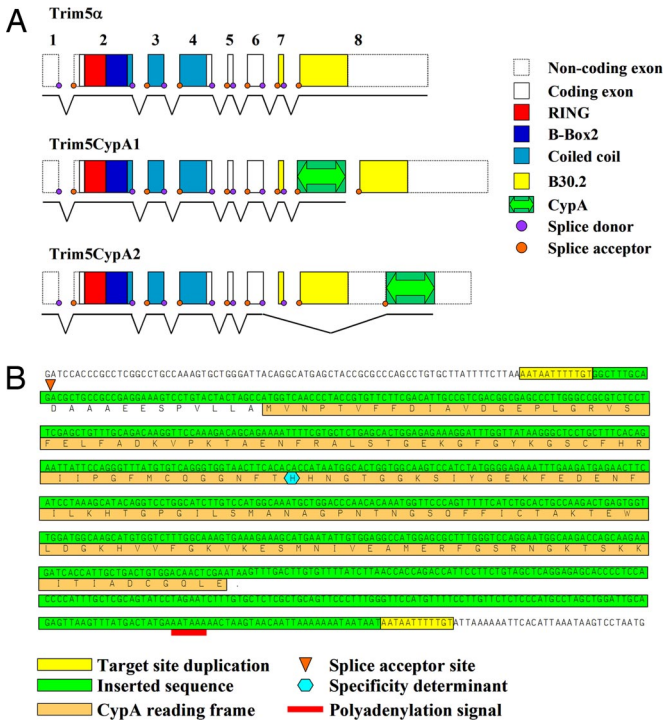

Fig. 1. Genomic organization and sequence of Trim5CypA2. (*A*) Organization of Trim5 gene showing the CypA insertions in Trim5CypA1 and Trim5CypA2. Coding exons (solid lines) and noncoding exons (dashed lines) are drawn to scale. Exon numbers are found at the top, and the mRNA derived from splicing are indicated below. The Trim5 domains (RING, B-Box, CC, B30.2, and CypA) are color-coded. Splice donors and acceptors are depicted as circles at the bottom of the boxes representing the exons. CypA is inserted between exons 7 and 8 in Trim5CypA1, where exon 7 is spliced to CypA instead of exon 8. In Trim5CypA2, CypA is inserted within the noncoding part of exon 8, and inactivation of the splice acceptor site in exon7 results in the splicing of exon 6 to CypA. (*B*) Sequence of the CypA insertion in Trim5CypA2. The inserted sequence, which is highlighted in green, is flanked by target site duplications, which are highlighted in yellow. The CypA reading frame is colored orange. The splice acceptor site (orange arrowhead), HIV-2 specificity determinant (cyan hexagon), and polyadenylation signals (red bar) are indicated.

A Second Fusion Event

This second chimeric gene, which we will refer to as Trim5CypA2, was identified in three different species of macaque monkeys, *Macaca mulatta* (rhesus macaque) (8, 11), *Macaca nemestrina* (pig-tailed macaque) (9–12), and *Macaca fascicularis* (long-tailed or crab-eating or cynomolgus macaque) (10). Like Trim5CypA1 from *Aotus*, Trim5CypA2 appears to have been generated by LINE element-mediated insertion (13) of a CypA pseudogene within the TRIM5 locus. However, the pseudogenes are found at two different locations; in Trim5CypA1 it lies between TRIM5 exons 7 and 8, whereas in Trim5CypA2 it is present in the untranslated region of exon 8 (Fig. 1*A*). It was present, in homozygous form, in >20 pig-tailed macaques and in both of the cynomolgus macaques examined. By contrast, it represents one of seven Trim5 alleles found in rhesus macaques (8, 14). Interestingly, it appears to be present in Indian (allele frequency 25%) but absent in Chinese rhesus macaques (8). The inheritance of Trim5CypA2 might prove a useful tool for examining unresolved issues regarding the evolution of different macaque species (15).

Trim5CypA1 is translated from mRNA formed by splicing from the end of TRIM5 exon 7 to a splice acceptor found in the CypA insertion just upstream of the CypA coding sequences, whereas the Trim5CypA2 mRNA is formed by joining the end of TRIM5 exon 6 to the CypA splice acceptor site (Fig. 1*A*). A SNP inactivating the exon 7 splice acceptor cosegregates with the retrotransposed CypA sequences in Trim5CypA2 (8, 11), presumably resulting in increased splicing to CypA. However, we cannot be certain whether this SNP occurred before or after CypA insertion.

Surprisingly, the Trim5CypA2 gene product, unlike that from Trim5CypA1,

Author contributions: J.P.S. and M.W.Y. wrote the paper.

The authors declare no conflict of interest.

See companion articles on pages 3557, 3563, and 3569.

*To whom correspondence should be addressed. E-mail: jstoye@nimr.mrc.ac.uk.

does not inhibit HIV-1 replication. It is therefore possible that macaques in which Trim5CypA2 has replaced TRIM5α might provide useful systems for studies involving HIV-1 growth. However, Trim5CypA2 clearly has restriction activity because it can block infection by HIV-2, simian immunodeficiency virus $(SIV)_{AGM}$tan, and feline immunodeficiency virus (FIV) (8, 9). This difference in restriction properties results from a point mutation leading to an H69R change in the CypA domain (Fig. 1*B*) (9). Like the Trim5CypA1 protein and Trim5α, the Trim5CypA2 gene product appears to block virus replication before reverse transcription of viral RNA in a proteasome-dependent fashion; as with the other factors, blocking proteasome function with MG132 leads to the nonproductive accumulation of nonintegrated viral DNA (8). Restriction of HIV-2 and SIV_{AGM}tan by the Trim5CypA2 gene product can be completely inhibited by treatment of cell cultures with the CypA binding drug, cyclosporine A, indicating a direct interaction between the Trim5CypA2 protein and CA (9) (8). By contrast, cyclosporine A only partially reverses Trim5CypA2 inhibition of FIV replication, most likely because the H69R change increases the affinity for the FIV CA protein.

Random or Predictable Evolution?

The description of the original Trim5CypA fusion was greeted with considerable surprise. Reports of a second case of exon shuffling involving the same gene parents prompt a closer examination of whether twice is just a coincidence or really part of a pattern that will be confirmed by the identification of further hybrids. It may be instructive to consider the example of retroviral vectors designed for gene finding. Basic features include a system for high-frequency insertion into chromosomes, an efficient splice acceptor system to join genes, a polyadenylation signal to stabilize transcripts, and a selectable/visible marker gene to detect gene splicing (16). CypA can provide all of these properties. Database mining reveals that CypA is one of the most common pseudogenes, implying that retrotransposition of this sequence can occur at high frequency (17). In addition, a splice acceptor signal is found just upstream of the CypA ORF, and a polyadenylation signal is found downstream (Fig. 1*B*). Further, resistance to retroviral infection could clearly provide a strong evolutionary advantage allowing selection of fusion genes incorporating CypA in a fashion leading to retrovirus restriction. Multimerizing CypA moieties by fusion to other Trim genes or the murine Fv1 gene *in vitro* has been shown to generate novel restriction factors (18, 19), so the choice of the upstream partner would appear more flexible. However, fusion to Trim5 may give a more efficient restriction factor perhaps because of proteasome recruitment. Alternatively, Trim5 may be truly dispensable for normal cell physiology, allowing gene fusions at little cost to the organism, unlike other members of the Trim family.

Although we have no way of knowing which retrovirus provided the selective forces in *Aotus* or *Macaca* for Trim5CypA1 or Trim5CypA2, selection certainly seems to have occurred, because all 10 species of owl monkey have Trim5CypA1 (20) and members of three of the major groups (15) of macaques have Trim5CypA2, thereby confirming the protective value of the Trim5CypA chimera. It thus appears abundantly clear that Trim5CypA chimeric proteins can block viral infections, but it remains to be seen what fraction of infections can be resisted in this manner. Might this represent the natural shape of things to come in the battle against AIDS?

1. Bieniasz PD (2004) *Nat Immunol* 5:1109–1115.
2. Stremlau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J (2004) *Nature* 427:848–853.
3. Wu X, Anderson JL, Campbell EM, Joseph AM, Hope TJ (2006) *Proc Natl Acad Sci USA* 103:7465–7470.
4. Meroni G, Diez-Roux G (2005) *BioEssays* 27:1147–1157.
5. Sayah DM, Sokolskaja E, Berthoux L, Luban J (2004) *Nature* 340:569–573.
6. Nisole S, Lynch C, Stoye JP, Yap MW (2004) *Proc Natl Acad Sci USA* 101:13324–13328.
7. Lin T-Y, Emerman M (2006) *Retrovirology* 3:70.
8. Wilson SJ, Webb BLJ, Ylinen LMJ, Verschoor E, Heeney JL, Towers GJ (2008) *Proc Natl Acad Sci USA* 105:3557–3562.
9. Virgen CA, Kratovac Z, Bieniasz PD, Hatziioannou T (2008) *Proc Natl Acad Sci USA* 105:3563–3568.
10. Brennan G, Kozyrev Y, Hu S-L (2008) *Proc Natl Acad Sci USA* 105:3569–3574.
11. Newman RM, Hall L, Kirmaier A, Pozzi L-A, Pery E, Farzan M, O'Neil S, Johnson W (2008) *PLoS Pathogens* 4(2):e1000003.
12. Liao C-H, Kuang Y-Q, Liu H-L, Zheng Y-T, Su B (2007) *AIDS* 21(Suppl 8):S19–S26.
13. Ostertag EM, Kazazian HH, Jr (2001) *Annu Rev Genet* 35:501–538.
14. Newman RM, Hall L, Connole M, Chen GL, Sato S, Yuste E, Diehl W, Hunter E, Kaur A, Miller GM, *et al*. (2006) *Proc Natl Acad Sci USA* 103:19134–19139.
15. Tosi AJ, Morales JC, Melnick DJ (2000) *Mol Phylogenet Evol* 17:133–144.
16. Skarnes WC (1990) *Bio/Technology* 8:827–831.
17. Zhang Z, Harrison PM, Liu Y, Gerstein M (2003) *Genome Res* 13:2541–2558.
18. Yap MW, Dodding MP, Stoye JP (2006) *J Virol* 80:4061–4067.
19. Yap MW, Mortuza GB, Taylor IA, Stoye JP (2007) *Virology* 365:302–314.
20. Ribeiro IP, Menezes AN, Moreira MA, Bonvicino CR, Seuanez HN, Soares MA (2005) *J Virol* 79:14998–15003.