

The Birth of a Viral Gene

September 16, 2013

G Brennan

Viruses are constrained by their small genome size, and have therefore evolved several mechanisms to maximize their coding potential. One such adaptation is a process called “overprinting” in which one gene is encoded on top of another gene in a different reading frame (see figure). However, it is unclear how the overprinted gene evolves without disrupting the pre-existing gene. In a recent study published in the Proceedings of the National Academy of Sciences, Drs. Joseph J. Carter, Matthew D. Daugherty, Harmit S. Malik and Denise A. Galloway (Human Biology, Basic Sciences, and Public Health Sciences Divisions), along with a team of collaborators, identify an overprinted gene in Merkel cell polyomavirus (MCPyV), and define conserved features that suggest a possible mechanism for the origin of this overprinted gene.

MCPyV, which was discovered in a rare but deadly form of skin cancer (Merkel cell carcinoma), contains an ~200 amino acid expanded and divergent region in the large T (LT) protein relative to other polyomaviruses. The early region (ER) of polyomaviruses, which includes LT, is a hotspot for generating protein diversity; therefore, the authors hypothesized that this divergent region of MCPyV may encode a novel protein. Surprisingly, sequence analysis of this region identified a predicted open reading frame (ORF) in the +1 frame relative to the second exon of LT. To determine if the overprinted ORF encoded a protein, the authors transfected cells with the wildtype MCPyV genome and identified a protein of the appropriate size from cells by immunoblot analysis. This protein was not detected in cells transfected with a genome containing a point mutation disrupting the overprinted ORF but not LT, confirming protein expression from this alternate reading frame. The authors named this new protein Alternate Large T Open reading frame (ALTO). ALTO was not necessary for DNA replication, so the authors suggest ALTO may play an accessory role in the viral lifecycle, similar to many other overprinting ORFs.

Although all polyomaviruses encode the overprinted LT protein, ALTO has not been described in other polyomaviruses. Therefore, the authors investigated the evolutionary origin of this new gene. By comparing DNA sequences among viruses, the researchers determined that ALTO is only present in hominid polyomaviruses. However, the authors noted that the murine polyomavirus (MPyV) Middle T (MT) gene is also overprinted on the second exon of LT, and ALTO and MT share a basic motif upstream of a hydrophobic C terminus. These similarities raised the possibility that ALTO and MT, despite having almost no sequence similarity, share a common ancestor. To

investigate this hypothesis, the authors searched for putative overprinting ORFs in all known polyomaviruses. This analysis identified one clade of polyomaviruses that all potentially encode an ALTO - like gene, while the polyomaviruses outside this lineage contain several stop codons within this reading frame (see figure). In many cases, this novel protein represents only the fifth or sixth known protein in these viruses, and thus defines a new clade of polyomaviruses that the authors name ALTO or middle T containing polyomaviruses (Almipolyomaviruses).

The ALTO/MT - like genes discovered in this study share very little sequence similarity; however, they share three conserved features 1) LT gene expansion, 2) a conserved YGS/T motif in LT that overlaps exactly with the ALTO start codon, and 3) a C - terminal hydrophobic domain. Interestingly, MPyV MT is an oncoprotein that is sufficient for cancer development in rodents. "Since we still do not know how Merkel cell polyomavirus causes cancer, it will be interesting to see whether the properties that allow the rodent virus proteins to cause cancer are also present in ALTO from Merkel cell polyomavirus," said Dr. Daugherty.

[Carter JJ, Daugherty MD, Qi X, Bheda-Malge A, Wipf GC, Robinson K, Roman A, Malik HS, Galloway DA.](#) 2013. Identification of an overprinting gene in Merkel cell polyomavirus provides evolutionary insight into the birth of viral genes. *Proc Natl Acad Sci U S A.* 110(31):12744 - 9.

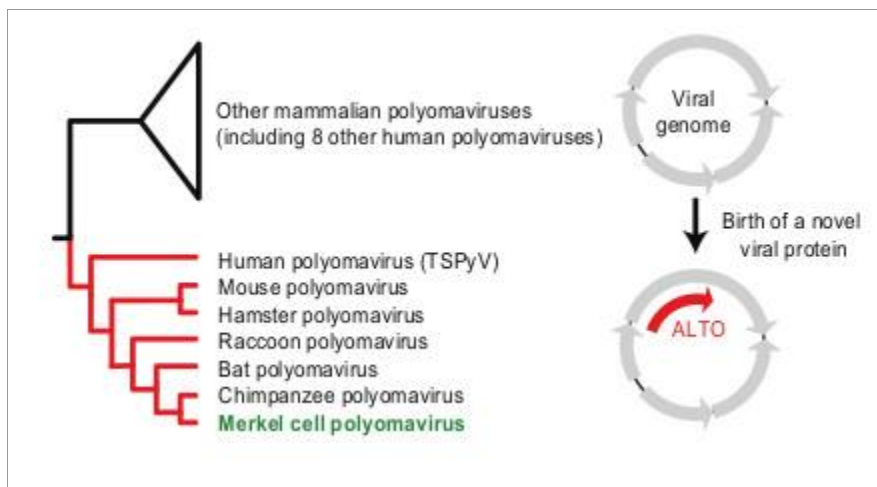


Image courtesy Dr. Matthew Daugherty

Phylogenetic tree of polyomaviruses, with red branches indicating viruses that contain ALTO - like proteins. The human oncovirus, Merkel cell polyomaviruses, in which ALTO was first discovered, is indicated in green. On the right is a schematic of the viral genome with grey arrows indicating proteins present in all polyomaviruses. The red arrow indicates the position of ALTO within the overlapping Large T antigen.