A Genetic Element Common to Tumor Viruses and Human

Angelo Kontgas[1] and Bino John[2].
[1]Chemistry and Biochemistry Department, Utah State university, Logan, UT 84322 USA
[2]Department of Computational Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15213 US

Abstract

Several viruses that predispose humans and animals to the development of cancer are known[1,2]. We hypothesized that such viruses may have common genomic signatures that help promote tumorigenesis in their hosts. Therefore, we investigated whether sequence elements that are conserved between humans and a set of 12 cancer-associated viruses and can be identified. The Mouse Mammary Tumor virus (MMTV), in stark contrast to all other 11 viruses displayed significant sequence similarity to the human genome. We identified 28 instances of similarity between MMTV and human genome. Four unique segments of MMTV DNA are incorporated at 28 locations in the human genome. One of the four MMTV segments is also evolutionarily preserved in three other cancer-associated viruses. Evolutionary analysis of the viral segments and the human DNA indicate that humans were originally infected by a variant of this virus through the consumption of animal milk.

Introduction

The past several decades of research have linked several viruses that predispose humans and animals to the development of cancer. The molecular mechanisms which lead to the development of cancers that are linked to cancer-associated viruses are not clearly understood. We endeavored to test whether cancer-associated viruses contain common genetic elements that are involved in cancer and are conserved over millions of years of evolution. Unexpectedly highly conserved genomic regions between distantly related species such as human, mouse, fish and flies have recently been identified[3]. Therefore, we investigated whether similar sequence-level comparisons of viral and human genomes can discover unusually conserved sequence elements.

Methods

A list of virus genomes associated with cancer was curated. The genomes were compared using BLAST[4] with each other and the human genome. The analysis yielded several sequences that were common ("hits") to viruses and human. The hits were then compared to a non-redundant set of all known protein sequences, the genomes of all known genomes of higher order organisms, the expressed sequence tags (EST) of known human genes, and viral genomes. Evolutionary relationships between related sequences were deduced based on multiple sequence alignments (CLUSTALW[5]) of related hits. To probe the possibility of highly stable DNA/RNA structures in conserved hits (eg; microRNA-like genes), candidate sequences were folded using MFOLD[6].

Results

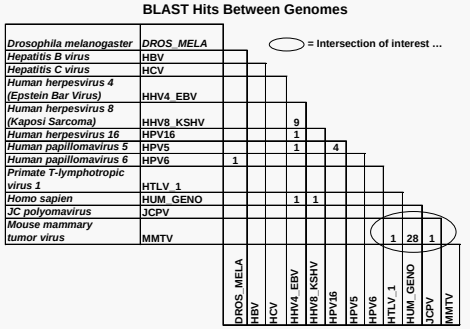

Figure 1. The matrix provides a summary of BLAST hits between various viruses and human genome. The circle represents the conserved sequences identified between MMTV, the Human Genome and two virus genomes, HTLV_1 and JCPV. Note that MMTV is 95% homologous with Human Mammary Tumor Virus (HMTV)[7].

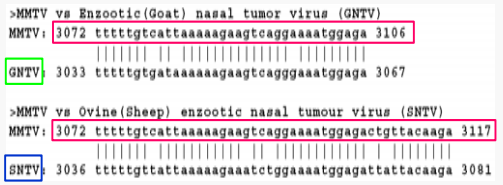

Figure 2. One of three sequences (MMTV-S1) conserved between MMTV and human genome is also preserved in Enzootic (Goat) nasal tumor virus (GNTV), and the Ovine (Sheep) nasal tumor virus (SNTV).

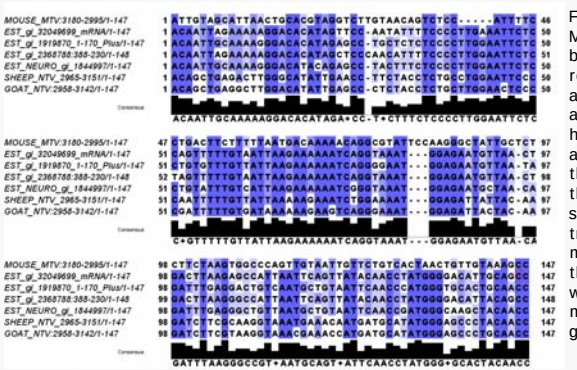

Figure 3.
Multiple sequence alignment between MMTV-S1 and its related hits in GNTV, SNTV and the human ESTs. The alignment indicates that the human genomic elements are considerably closer to the sheep and goat NTVs than to MMTV. Retroviruses such as NTVs are known to transfer between species via milk. Our study suggests that a viral relative of NTV was passed onto humans via milk from an infected sheep, goat or their relatives.

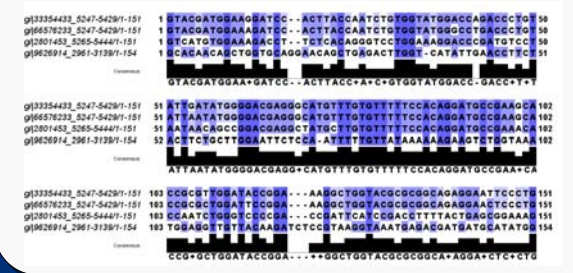

Figure 4.
A CLUSTALW alignment suggests extensive similarity between the GNTV and SNTV. less similarity with MMTV, and even less similarity with Sheep adenocarcinoma virus.

Results

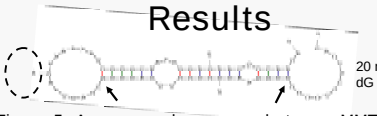

Figure 5. A conserved sequence between MMTV and a human expressed sequence tag folds into a stem loop structure that is characteristic of microRNA-like genes.

Conclusion

HMTV EVOLUTION ?

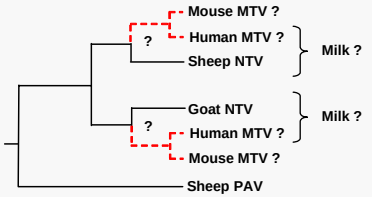

Figure 6. The 95% homology of HMTV with MMTV suggests a common ancestor with either SNTV or GNTV. It also suggests the mechanism of passage to human is through the animal's milk.

Future Work

- What is the functional significance of the 28 hits of MMTV on the 18 human chromosomes?
- What is the significance of the evolutionarily preserved sequences between human genome, MMTV, SNTV, and GNTV ?

Acknowledgements

The national BBSI program (http://bbsi.eeicom.com) is a joint initiative of the NIH-NIBIB and NSF-EEC, and the BBSI @ Pitt is supported by the National Science Foundation under Grant EEC-0234002.
- Dr. Bino John
- Dept Computational Biology, U. Pitt

References

1. McCance, D. J. *Human tumor viruses*. American Society for Microbiology, Washington, D.C (1998).
2. Pellicano, R., Mladenova, I., Martinotti, R., Fagoonee, S. & Rizzetto, M. [Gastric cancer and Helicobacter pylori: an interdisciplinary point of view]. *Minerva Med.* **97**, 31-38 (2006).
3. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034-1050 (2005).
4. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402 (1997).
5. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680 (1994).
6. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406-3415 (2003).
7. Liu, B. *et al.* Identification of a proviral structure in human breast cancer. *Cancer Res.* **61**, 1754-1759 (2001).