Transcriptional Regulatory Elements and Transcription Factors that Control Kaposi's Sarcoma Human Virus Genes.

Manjula Kasoji[1,2] ,Dr. Bino John[1]

[1]*Bioengineering & Bioinformatics Summer Institute, Dept. of Computational Biology, University of Pittsburgh, 15260*

[2] *Bioinformatics Department, Georgia Institute of Technology, 30332*

Abstract

•**Kaposi's sarcoma human virus (KSHV)** is a tumor developing agent
•Its genes interfere with tumor suppressor pathways
and modify the host cellular environment
•**Goal:** identify the transcriptional sites and factors that are involved in regulating the KSHV genome.
•In reaching this goal we have:
-characterized the upstream regions of all the protein coding genes.
-These upstream sequences were loaded into a web program, **CLOVER** to identify **transcription factors**
•With the identification of these transcription factors we hope to better understand the regulation and infection of the Kaposi virus.

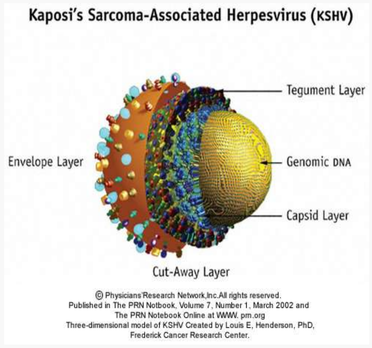
Introduction

Kaposi's sarcoma is now the fourth most common cancer caused by an infectious agent worldwide following gastric, cervical, and hepatic cancers. It is a cancerous disease that causes abnormal tissue growth under the skin and leads to lesions. KSHV is usually not life threatening unless it reaches to the internal organs such as lungs, liver, and gastrointestinal tract. Finding out more about transcriptional regulation may allow us to help decrease the spread of KSHV by disrupting its function. We believe that the identification of regulatory elements and factors is an attainable goal because KSHV uses human transcription factors to carry out its function.

We will attempt to identify regulation sites or motifs by using a computational method called K-factor. K-Factor predicts regulatory motifs in a set of functionally related sequences [Li]. KSHV is not evolutionary conserved therefore we are using K-factor because it does not rely on evolutionary conservation, and it is not based on existing motifs from a database as opposed to other motif prediction programs. The transcription factors will also be identified by using other computational programs that predict human transcription factors. KSHV does not have transcription factors of its own. It uses human factors to regulate its transcription. This is why we are searching for associated human transcription factors.

Method

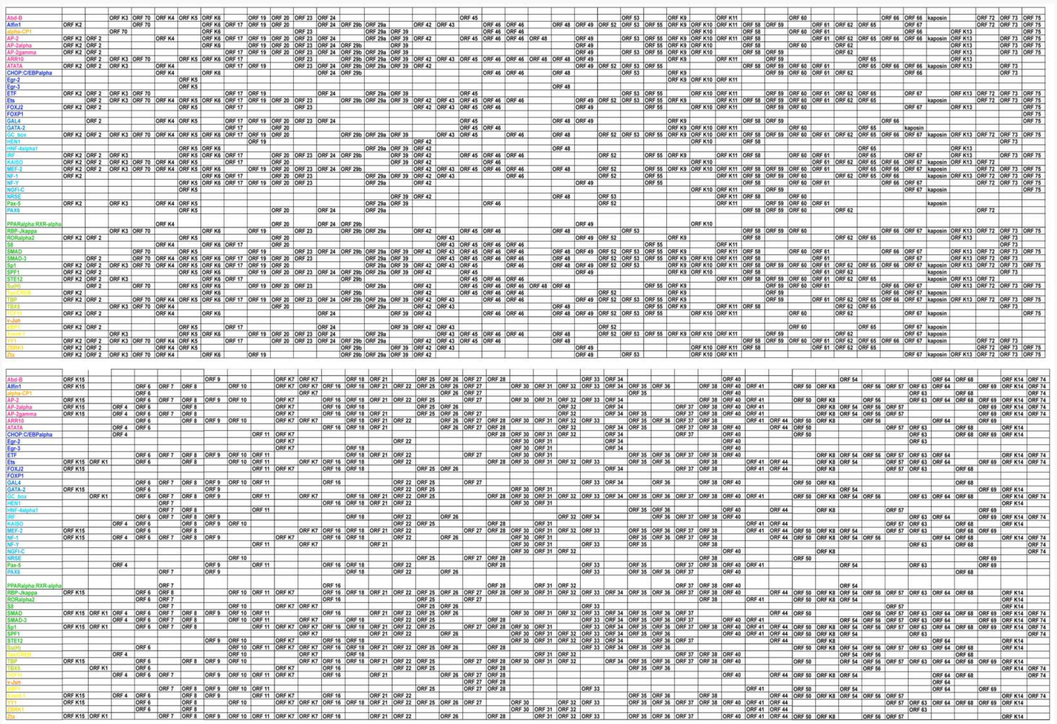
1. **Identification of transcription factors and motifs**
   - obtaining the KSHV genome
   - identifying all of the protein encoding genes
   - finding the upstream regions of up to one thousand nucleotides of the protein coding genes
   - using programs to identify the factors and motifs that are involved in these upstream regions
2. **Validation of these findings**
   - random sequences will be generated by a simple Perl script.
   - these random sequences will act as reference or control sequences to which the upstream sequences will be compared against.



Mortiz Kaposi

Kaposi's sarcoma on the skin of an AIDS patient.

Results



•Transcription factors and respective regulated genes.
•With the identification of these transcription factors we hope to better understand the regulation and infection of the Kaposi virus.

Future Research

• Find out which transcription factors regulate miRNAs

-miRNAs are a class of endogenous, small RNAs that are thought to negatively regulate protein production.

-aberrant expression of miRNAs is linked to cancer and other diseases

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.

1) Dr. Bino John
2) University of Pittsburgh

References

Bulyk, M. L. "Computational prediction of transcription-factor binding site locations." Genome Biol. 5.1 (2003): 201.

Fukao, T., et al. "An evolutionarily conserved mechanism for microRNA-223 expression revealed by microRNA gene profiling." Cell 129.3 (2007): 617-31.

Li, Ji., et al. "Regulatory Circuit of Human MicroRNA Biogenesis." PLoS Computational Biol. 3 (2007): 0721-0732.