



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



Manjula Kasoji^{1,2}, Dr. Bino John¹

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

²Bioinformatics 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

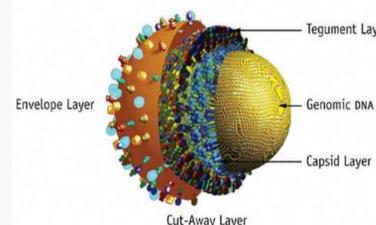
1. Identification of transcription factors and motifs

- 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.

Kaposi's Sarcoma-Associated Herpesvirus (KSHV)



© Physicians Research Network, Inc. All rights reserved. Published in The PRN Notebook, Volume 7, Number 1, March 2002 and The PRN Notebook Online at WWW.PRN.org Three-dimensional model of KSHV Created by Louis E. Henderson, PhD, Frederick Cancer Research Center.



Mortiz Kaposi

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

Results

Gene	Transcription Factor	Score	Position
ORF 1	TF1	0.85	100-200
ORF 2	TF2	0.72	150-250
ORF 3	TF3	0.91	120-220
ORF 4	TF4	0.68	180-280
ORF 5	TF5	0.79	140-240
ORF 6	TF6	0.83	160-260
ORF 7	TF7	0.75	130-230
ORF 8	TF8	0.88	170-270
ORF 9	TF9	0.71	110-210
ORF 10	TF10	0.86	190-290
ORF 11	TF11	0.73	155-255
ORF 12	TF12	0.81	125-225
ORF 13	TF13	0.69	185-285
ORF 14	TF14	0.77	145-245
ORF 15	TF15	0.84	165-265
ORF 16	TF16	0.72	135-235
ORF 17	TF17	0.89	175-275
ORF 18	TF18	0.74	115-215
ORF 19	TF19	0.87	195-295
ORF 20	TF20	0.70	150-250
ORF 21	TF21	0.82	130-230
ORF 22	TF22	0.76	170-270
ORF 23	TF23	0.80	140-240
ORF 24	TF24	0.73	180-280
ORF 25	TF25	0.85	160-260
ORF 26	TF26	0.71	120-220
ORF 27	TF27	0.88	200-300
ORF 28	TF28	0.75	160-260
ORF 29	TF29	0.83	140-240
ORF 30	TF30	0.77	180-280
ORF 31	TF31	0.81	155-255
ORF 32	TF32	0.69	125-225
ORF 33	TF33	0.86	205-305
ORF 34	TF34	0.72	165-265
ORF 35	TF35	0.84	145-245
ORF 36	TF36	0.78	185-285
ORF 37	TF37	0.82	150-250
ORF 38	TF38	0.70	130-230
ORF 39	TF39	0.89	210-310
ORF 40	TF40	0.74	170-270
ORF 41	TF41	0.87	150-250
ORF 42	TF42	0.80	130-230
ORF 43	TF43	0.76	170-270
ORF 44	TF44	0.85	150-250
ORF 45	TF45	0.71	130-230
ORF 46	TF46	0.88	215-315
ORF 47	TF47	0.75	175-275
ORF 48	TF48	0.83	155-255
ORF 49	TF49	0.77	195-295
ORF 50	TF50	0.81	165-265
ORF 51	TF51	0.72	135-235
ORF 52	TF52	0.89	220-320
ORF 53	TF53	0.74	180-280
ORF 54	TF54	0.87	160-260
ORF 55	TF55	0.80	140-240
ORF 56	TF56	0.76	180-280
ORF 57	TF57	0.85	160-260
ORF 58	TF58	0.71	140-240
ORF 59	TF59	0.88	225-325
ORF 60	TF60	0.75	185-285
ORF 61	TF61	0.83	165-265
ORF 62	TF62	0.77	205-305
ORF 63	TF63	0.81	175-275
ORF 64	TF64	0.72	145-245
ORF 65	TF65	0.89	230-330
ORF 66	TF66	0.74	190-290
ORF 67	TF67	0.87	170-270
ORF 68	TF68	0.80	150-250
ORF 69	TF69	0.76	190-290
ORF 70	TF70	0.85	170-270
ORF 71	TF71	0.71	150-250
ORF 72	TF72	0.88	235-335
ORF 73	TF73	0.75	195-295
ORF 74	TF74	0.83	175-275
ORF 75	TF75	0.77	215-315
ORF 76	TF76	0.81	185-285
ORF 77	TF77	0.72	155-255
ORF 78	TF78	0.89	240-340
ORF 79	TF79	0.74	200-300
ORF 80	TF80	0.87	180-280
ORF 81	TF81	0.80	160-260
ORF 82	TF82	0.76	200-300
ORF 83	TF83	0.85	180-280
ORF 84	TF84	0.71	160-260
ORF 85	TF85	0.88	245-345
ORF 86	TF86	0.75	205-305
ORF 87	TF87	0.83	185-285
ORF 88	TF88	0.77	225-325
ORF 89	TF89	0.81	195-295
ORF 90	TF90	0.72	165-265
ORF 91	TF91	0.89	250-350
ORF 92	TF92	0.74	210-310
ORF 93	TF93	0.87	190-290
ORF 94	TF94	0.80	170-270
ORF 95	TF95	0.76	210-310
ORF 96	TF96	0.85	190-290
ORF 97	TF97	0.71	170-270
ORF 98	TF98	0.88	255-355
ORF 99	TF99	0.75	215-315
ORF 100	TF100	0.83	195-295

• 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.