

## VIRUS GENOME IMAGING VIA a<sup>2</sup>GRAMS: BUILDING A MATLAB TOOLBOX FOR PROKARYOTIC DNA ANALYSIS

*C. Dantas Silva*<sup>1</sup>; *E.A. Bouton*<sup>1</sup>; *N.S. Santos-Magalhães*<sup>2</sup>; *H.M. de Oliveira*<sup>1</sup>

Universidade Federal de Pernambuco, <sup>1</sup>Departamento de Eletrônica e Sistemas; <sup>2</sup>Departamento de Bioquímica, Laboratório de Imunologia Keizo Asami, LIKA

Much of genomic signal analysis approaches for feature extraction and functional cataloguing have been focused on oligonucleotide patterns in the linear primary sequences of genomes. New DNA-imaging tools for genomic signal processing namely codongrams and a<sup>2</sup>grams had recently been offered for extracting meaningful genomic features embedded in DNA. A Matlab<sup>TM</sup> toolbox was implemented for allowing the image analysis of viruses and bacteriophages. Twenty different a<sup>2</sup>grams are defined for a genome, one for each amino acid (valgram is an a<sup>2</sup>gram for valine; alagram is an a<sup>2</sup>gram for alanine, etc.) They furnish information about the distribution and occurrence of the investigated amino acid. The codongram describes the distribution of a specific codon through the genome. The a<sup>2</sup>gram for a particular amino acid provides information about the sections of the DNA strand, which potentially leads to the synthesis of such an amino acid. DNA ×grams are among powerful visual tools for GSA like spectrograms, which can be applied when searching for particular nucleotide patterns. Among such patterns, the software includes built-in options the following: metgram to find out potential start position of genes, Shine-Dalgarno sequence localizer (translation mRNA → protein), TATA Box (replication DNA → mRNA), Enter a sequence (DNA particular sequence finder). A few genomes of viruses and bacteriophage were made available in the DEMO version: Bacteriophage ΦX174, phage MS2, Tomato Bushy Stunt Virus (TBSV), Tobacco Mosaic Virus (TMV), Phage M13, and Simian virus SV40 (genome lengths ranging from 3,569 to 6,400 bp). This tool is particularly helpful for comparing viruses, and it is also particularly valuable for educational purposes.