

Analysis and classification of long terminal repeat sequences from plant LTR-retrotransposons

Author: Jakub Horváth^{1*}

Supervisor: Matej Lexa^{1†}

Faculty of Informatics, Masaryk University, Botanická 68A,
602 00 Brno-Královo Pole, Czech Republic

Long Terminal Repeats (LTRs) are DNA sequences common in various organisms. By understanding their structure we may be able to understand the evolutionary processes of the organisms in which they are located.

These sequences are, however, difficult to detect using common methods in bioinformatics

This thesis aims to analyze their inner structure, as well as train machine learning models which are able to recognize LTR sequences. Different models of different complexities are used for this approach in order to find the best one.

The findings of this work may be used by scientists studying evolution of viruses in plants.

*jakubhorvath119@gmail.com

†lexa@fi.muni.cz