Abstract

This diploma thesis deals with a development of a strategy for data evaluation generated by next-generation sequencing. Using bioinformatics tools such as Galaxy, Velvet and Enterovirus genotyping tool new aproach of data processing was optimized. There were 22 samples analyzed which of 10 were grown on cell culture. Remaining 12 were obtained from real stool samples. All samples were taken from children at the highest genetic risk of type 1 diabetes. All of them were enterovirus positive. Enteroviruses and their following infections have been suspecting to be involved in ehiology of type 1 diabetes for a long time. That's a disease resulting to an absolut insulin deficiency due to autoimmune destruction of pancreatic beta cells. Genetic components seems to be relatively well defined (the HLA, INS, STLA4, PTPN22, CTLA4, IFIH1 and numerous other genes), the environmental part of the etiology remains obscured.

We were able to assemble 22 genomes de novo. However, there were numerous gaps among the particular contigs. For the first nine samples these gaps were complemented by Sanger sequencing. Nine full-length genomes were assempled this way.

The main contribution of this work was to create a universal process of analyzing data from next-generation sequencing. This has already been using for further analysis of the samples which are subjected to this type of sequencing. Using this procedure we are able to identify viruses from a sample without their specific detection.