Abstract: One of the most fundamental tasks in the analysis of biological sequences is the search for similarity. The immense amount of sequence data that has become available within the last two decades makes this search a challenging task - highly efficient but nevertheless accurate approaches are sought.

In the first part, I will discuss the SWIFT filter algorithm [1], a very efficient q-gram counting algorithm for finding high similarity regions in long sequences. The algorithm stands out from other efficient similarity search approaches by providing a guarantee for full sensitivity. This guarantee is achieved through elaborate calculation of algorithm parameters. At the end of this part, I will touch on the counterintuitive effects that the parameter calculation has on the filtering performance.

In the second part, I will briefly summarize the general strategy for computing multiple whole-genome alignments and cover the main contributions of my thesis to three aspects of this field: local alignment, representation in graphs, and genome rearrangement.