In this thesis we explore and compare various methods for efficient k-mer set representation. We evaluate traditional de Bruijn graph representation techniques against greedy approximation algorithms for the Shortest Superstring Problem. We describe the linear- time implementation of the well-known Greedy algorithm by Ukkonen [1990] and extend it to another related algorithm, called TGreedy. In addition, we test selected algorithms on a bacterial genome and pangenome to highlight the differences in the size of their output representation and the computational resources used, providing an insight into their respective efficiencies.