DNA read mapping is an important problem in Bioinformatics. With the introduction of next-generation sequencing (NGS) technologies, we are facing an exponential increase in the amount of genomic sequence data. The success of many medical and genetic applications critically depends on computational methods to process the enormous amount of sequence data quickly and accurately. However, due to the repetitive nature of human genome and limitations of the sequencing technology, current read mapping methods still fall short from achieving both high performance and high sensitivity.

A major challenge in building fast and sensitive mappers is that to increases sensitivity, a mapper has to increase the search space for mappings which inevitably generates false mappings in the process. False mappings incur unnecessary computation as false mappings are excluded from the final report. In this dissertation, we provide novel computational techniques to mitigate the impact of unnecessary computation generated by false mappings, while maintaining high sensitivity. Specifically, we propose 1) a novel SIMD-friendly bit-parallel filtering problem that quickly estimates false mappings without spending much computational resources; 2) a generalization of an efficient, state-of-the-art approximate string matching algorithm to support a wider range of penalty metrics; 3) a novel seed selection algorithms that reduces the number of false mappings without compromising sensitivity and 4) a novel seeding concept that allows increasing mapping sensitivity without using more seeds. We show that all of our techniques effectively reduce the unnecessary computation associated with false mappings and together they build a foundation for future mappers.