Analysis of DNA sequence data to identify mutations in tumours

Anna Larsson[1], **Annica Wilzen**[2], **Frida Abel**[2], **Erik Kristiansson**[1]

[1] Mathematical Sciences, Chalmers University of Technology, Sweden
[2] Department of Clinical Genetics, Sahlgrenska Academy, Gothenburg University, Sweden

Exome sequencing is used to analyse the DNA sequences of the full set of genes in one individual using high-throughput sequencing. Each gene is typically covered multiple times by sequencing millions of short ($\sim$ 75 base pairs) DNA fragments. The generated DNA sequencing data contains however both random noise and systematic errors which make the computational and statistical analyses non-trivial. We have applied exome sequencing on pheochromacytoma, a hormone-producing tumour in the adrenal medulla. The aim of the study was to find tumour-specific mutations and therefore both tumour and normal tissue from each patient were sequenced and analysed in a pair-wise design. Since the frequency of reads that harbour mutations often is low due to tumour heterogeneity, a main challenge was to differentiate between true mutations and sequencing errors. We applied a method called Mutect, which uses a Bayesian classifier designed to be able to capture low-frequency mutations, followed by different filters to enhance specificity. By analysing tumour-normal pairs from 4 patients with benign disease and 5 patients with malignant disease, we found in total 172 somatic mutations. The average number of mutations was significantly higher in malignant tumours compared to benign ones. On this poster we will discuss the technical challenges in the data analysis and the biological results.

Keywords: DNA sequence analysis, cancer, mutations