Project: Parsing of next generation sequencing (NGS) data

NGS data are high throughput, with millions of reads being assigned in every run for each sample. The outputs from NGS runs can identify hundreds of variants for each sample when compared to a wild type reference sequence. For each variant, there are multiple quality parameters and biological findings that must be captured and stored in a format that support downstream analysis, including the association of a particular variant with a clinical outcome. We have built a preliminary tool using Excel Macros for the extraction and presentation of NGS data from the Ion Torren PGM platform. This project involves improving the sophistication of this tool to allow additional flexibility, and to support linkage of the NGS data to other data types. This would require development of an interface for data representation and data management as well as ability to communicate with other databases.

Contact Dr. Janice Glasgow (janice@cs.queensu.ca) and
Dr. Kathrin Tryshkin (tyryshki@cs.queensu.ca)