Carsten Wiuf
Bioinformatics Research Centre
e-mail: wiuf@cs.au.dk

Stochastic Modeling and Analysis of DNA Sequence Data from Heterogeneous Tumors

Many cancers are believed to have clonal origin, starting from a single cell with a defining mutation and further acquiring one or more additional mutations before the first cancerous cell is established. For example, in *Follicular Lymphoma*, a blood cancer, the total number of required mutations M is believed to be two of which the first is a translocation called t(14;18).

A population of cancer cells evolves further over time and accumulates genetic changes, many of which are random and others potentially beneficial for the cancer. Consequently, cells in different parts of a tumor might show differences in their genomes, or DNA. This phenomenon is referred to as genetic tumor heterogeneity and is comparable to the genetic heterogeneity observed in individuals in a population.

Here, I address the problem of modeling how the tumor evolves over time and accumulates changes in the DNA, starting from the initial cell with the defining mutation. The model is stochastic and relies on birth-death processes; it allows the first required M mutations to be under selective pressure, while the subsequent mutations are neutral. I show that there is a simple description of how the (stochastic) number of tumor cells in the system changes over time and that the model imposes constraints on parameters that determine the reproducibility and the survival of cells; thus the model leads to biological insight.

Further, the model leads to a simple way of simulating tumor evolution. Based on this, I show how a sample of DNA sequences taken from distinct parts of a heterogeneous tumor might be used to draw inference on model parameters and date the origin of the tumor, as well as the defining and subsequent mutations. The latter might have clinical importance as it provides an estimate of the time from tumor initiation to diagnosis.

Finally, I show a simple application to DNA sequence data from *Follicular Lymphoma* patients and outlining some further mathematical and statistical work to be done.