UNIVERSITY OF GOTHENBURG

Integrative genomic and survival analysis of breast tumors

AKADEMISK AVHANDLING

som för avläggande av medicine doktorsexamen vid Sahlgrenska Akademin vid Göteborgs Universitet offentligen försvaras i Arvid Carlsson salen, Academicum, Medicinaregatan 3

onsdagen den 5 december kl. 09.00

av

Szilárd Nemes

Fakultetsopponent
Assoc. Prof Dr. Markus Rignér
Faculty of Medicine, Lund University, Sweden

Avhandlingen baseras på följande delarbeten:

I. Nemes Sz., Parris T, Danielsson A, Kannius-Janson M, Jonasson JM, Steineck G, Helou K. Segmented regression, a versatile tool to analyze mRNA levels in relation to DNA copy number aberrations. *Genes, Chromosomes and Cancer*. 2012, 51(1): 77-82.

II. Nemes Sz, Parris TP, Danielsson A, Einbeigi Z, Steineck G, Jonasson JM, and Helou K. Integrative genomics with mediation analysis in a survival context. (*Submitted*)

III. Nemes Sz, Parris TP, Danielsson A, Jonasson JM, Genell A, Karlsson P, Steineck G and Helou K. A novel 12-gene panel predicting clinical outcome of breast cancer. (*Submitted*)

IV. Nemes Sz, Danielsson A, Parris TP, Jonasson JM, Karlsson P, Steineck G and Helou K. Permutation test for the clonal origins of multiple tumors. (*Manuscript*)

ABSTRACT

With the continued accumulation of genomic data at ever increasing resolution, the challenge ahead lies in reading out meaningful clinical/biological information from the data that can contribute to a better understanding of the cancerous process. The need for novel approaches and new statistical methods is therefore strong.

The present thesis aims to contribute to the field with three problem-specific applications that hopefully will aid researchers in a better understanding of genomic data.

The first paper exemplifies the adaptation of a piecewise-linear regression framework for integrative analysis of DNA copy number aberrations and gene expression (mRNA) data. The method helps to identify the association between copy number and gene expression, but it takes a further step and allows detection of changing patterns and changepoints that could serve as a proxy for the degree of genomic instability that causes disruptions in feedback-mechanisms.

The second paper advocates the adaptation of a mediation analysis for a concomitant analysis of DNA copy number aberrations, mRNA and survival data. The paper offers ways of statistical inference by means of the Delta method applicable concomitantly on a large number of genes. If a mediation effect is observed for a specific gene, we hypothesize that the specific gene is a driver gene. If no mediation effect is observed, possible associations between DNA copy number aberrations and the outcome are likely to indicate passenger genes.

The third paper is a more applied/clinical work using applied statistics which identified a novel panel of 12-genes that can serve as a prognostic tool for breast cancer specific survival.

The thesis concludes with a methodological description in which we describe an easy permutation-based approach for testing the clonal origins of multiple tumors. The main assumption of the proposed method is that if two tumors that share a common origin, or if the alleged secondary tumor is clonally related to the primary tumor, they share a higher and tumor-specific amount of matching chromosomal aberrations (gains or deletions) than recurrent chromosomal aberrations can explain.

Keywords: DNA copy number aberrations, messenger-RNA, breast cancer, regression, survival analysis, mediation, permutations

ISBN: 978-91-628-8538-0