Göteborg, 2018

Comprehensive analysis of structural genomic alterations in cancer

Akademisk avhandling

Som för avläggande av medicine doktorsexamen vid Sahlgrenska akademin, Göteborgs universitet kommer att offentligen försvaras i hörsal **Arvid Carlsson**, Academicum, Medicinaregatan 3, Göteborg

Fredagen den **2 mars** 2018, klockan **9:00**

av Babak Alaei-Mahabadi

Fakultetsopponent:

Dr. Giovanni Ciriello

University of Lausanne, Lausanne, Swizterland

Avhandlingen baseras på följande delarbeten

I. The landscape of viral expression and host gene fusion and adaptation in human cancer.
Tang KW, **Alaei-Mahabadi B**, Samuelsson T, Lindh M, Larsson E.
Nature Commun. 2013;4:2513.

II. Global analysis of somatic structural genomic alterations and their impact on gene expression in diverse human cancers.
Alaei-Mahabadi B, Bhadury J, Karlsson JW, Nilsson JA, Larsson E.
Proc Natl Acad Sci U S A (PNAS) . 2016;113(48):13768-13773.

III. Systematic investigation of promoter substitutions resulting from somatic intrachromosomal structural alterations in diverse human cancers.
Alaei-Mahabadi B, Larsson E.
Manuscript

SAHLGRENSKA AKADEMIN
INSTITUTIONEN FÖR BIOMEDICIN

GÖTEBORGS UNIVERSITET

Göteborg, 2018

Comprehensive analysis of structural genomic alterations in cancer

Babak Alaei-Mahabadi

Medicinsk kemi och cellbiologi, institutionen för biomedicin, Sahlgrenska akademin, Göteborgs universitet, Sverige, 2018.

Abstract

The transformation of a normal cell into a cancer cell involves the accumulation of somatic DNA alterations that confer growth and survival advantages. These genomic alterations can be different in terms of pattern and size, comprising single nucleotide variants (SNVs), small insertions or deletions (indels), structural variations (SVs) or foreign DNA insertions such as viral DNA. Cancer genomes typically harbor numerous such changes, of which only small fractions are driver events that are positively selected for during the evolution of the tumor. High throughput sequencing has enabled systematic mapping of somatic DNA alterations across thousands of tumor genomes. Mutations in particular have been thoroughly explored in this type of data, and this has implicated many new genes in tumor development. However, our knowledge remains more limited when it comes to the contribution of SVs to cancer. In the present thesis, we made use of publicly available cancer genomics data to gain further insight into the role of structural genomic alterations in tumor development.

Viruses cause 10-15% of all human cancers through multiple mechanisms, one of which is structural genomic changes due to viral DNA being integrated into the human genome. Thus, in the first study, we performed an unbiased screen for viral genomic integrations into cancer genomes. We developed a computational pipeline using RNA-Seq data from ~4500 tumors across 19 different cancer types to detect viral integrations. We found that recurrent events typically involved known cancer genes, and were associated with altered gene expression.

SVs can lead to copy number amplification of specific cancer driver genes, as well as the formation of fusion oncogenes, but their importance in cancer beyond these types of events is underexplored. We mapped SVs to the human genome using whole genome sequencing data from 600 tumors across 18 different cancer types and investigated the global relationship between SVs and mRNA changes. We found that such events often contribute to altered gene expression in human tumors, but we were not able to detect novel recurrent driver events. To increase the cohort size, we used a larger but lower resolution and more limited dataset, comprising of microarray based DNA copy number profiles from ~10,000 tumors across 32 cancer types, with the aim of identifying recurrent SV driver events in tumors. Specifically, we investigated SVs predicted to result in promoter substitution events, a known mechanism for gene activation in cancer, and found several recurrent activating events with potential cancer driver roles. Notable among our findings in all the studies were human papillomavirus integrations in *RAD51B* and *ERBB2* and gene fusions involving *NFE2L2, TIAM2* and *SCARB1,* all being known cancer genes.

Taken together, massive amounts of genomic and transcriptomic sequencing data allowed us to comprehensively map viral integrations and structural variations in cancer, which led to the identification of several genes with potential roles in tumor development.

Keywords: Somatic structural variations, viral integrations, gene fusions
ISBN 978-91-629-0422-7 (PRINT) http://hdl.handle.net/2077/54530
ISBN 978-91-629-0423-4 (PDF)