Bioinformatics

- Oxford Journals
- Life Sciences
- Bioinformatics
- Bioinformatics Advance Access
- 10.1093/bioinformatics/btn034

Bioinformatics Advance Access published online on February 8, 2008
Bioinformatics, doi:10.1093/bioinformatics/btn034

Integrative analysis reveals the direct and indirect interactions between DNA copy number aberrations and gene expression changes

Hyunju Lee [1,4,*], Sek Won Kong [2,3,*] and Peter J Park [1,2,†]

[1]Harvard-Partners Center for Genetics and Genomics, Boston, MA, USA
[2]Informatics Program, Children's Hospital, Boston, MA, USA
[3]Department of Cardiology, Children's Hospital, Boston, MA, USA
[4]Department of Information and Communication, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea

†To whom correspondence should be addressed. Prof. Peter J Park, E-mail: peter_park@harvard.edu

Abstract

Motivation: DNA copy number aberrations (CNAs) and gene expression (GE) changes provide valuable information for studying chromosomal instability and its consequences in cancer. While it is clear that the structural aberrations and the transcript levels are intertwined, their relationship is more complex and subtle than initially suspected. Most studies so far have focused on how a CNA affects the expression levels of those genes contained within that CNA.

Results: To better understand the impact of CNAs on expression, we investigated the correlation of each CNA to all other genes in the genome. The correlations are computed over multiple patients that have both expression and copy number measurements in brain, bladder, and breast cancer data sets. We find that a CNA has a direct impact on the gene amplified or deleted, but it also has a broad, indirect impact elsewhere. To identify a set of CNAs that is coordinately associated with the

expression changes of a set of genes, we used a biclustering algorithm on the correlation matrix. for each of the three cancer types examined, the aberrations in several loci are associated with cancer-type specific biological pathways that have been described in the literature: CNAs of chromosome (chr) 7p13 were significantly correlated with epidermal growth factor receptor signaling pathway in glioblastoma multiforme, chr 13q with NF-kappaB cascades in bladder cancer, and chr 11p with Reck pathway in breast cancer. In all three data sets, gene sets related to cell cycle/division such as M phase, DNA replication, and cell division were also associated with CNAs. Our results suggest that CNAs are both directly and indirectly correlated with changes in expression and that it is beneficial to examine the indirect effects of CNAs.

Availability: The code is available upon request.

Contact: peter_park@harvard.edu

Associate Editor: Dr. Chris Stoeckert

*these authors contributed equally to this work

Received on September 20, 2007; revised on December 29, 2007; accepted on January 22, 2008

Online ISSN 1460-2059 - Print ISSN 1367-4803