

CHESS (CgHExpreSS) : A software for association study of copy number changes and correlation analysis between copy number changes and gene expression

Mi-Kyung, Lee¹
mkLee77@khu.ac.kr

Yang-Seok, Kim¹
yskim1158@khu.ac.kr

¹ Department of Physiology, College of Oriental Medicine, KyungHee University, #1Hoegi-Dong, Dongdaemun-gu, Seoul 130-701, South Korea

Abstract

CHESS(CgHExpreSS) is a program for investigating the relationship of genetic alterations with both specific phenotype and gene expression. The association study module allows to detect regions of recurrent genomic alteration in each samples by analyzing relationship between genetic variations and a particular phenotype. The correlation analysis module elucidates the effects of genomic alteration on the level of gene expression. The visualization module of CHESS will help understand the analysis result intuitively and conveniently. As CHESS was developed on java application platform, it can be run under any operating system.

Keywords: CGH, microarray, gene expression, association study

1 Introduction

Recently, Comparative Genomic Hybridization (CGH) has emerged as a technique which detects copy number changes (gains/losses) in the DNA content of an individual. Especially, it is well known that copy number changes frequently occur in cancer patients, many researchers have attempted to take advantage of CNVs as the genetic marker for a particular disease. So, several programs have recently been introduced for analyzing CGH array data, such as CGH-Explorer, SeeGH, CGHAnalyzer and aCGHViewer[1,2,3]. However, these programs are focused on calling and visualizing copy number and do not provide information on phenotype specific CNVs which frequently occur in patients compared to normal samples. Another challenging point of CNVs is to find out the effects of genomic alteration on the level of gene expression, however, the tool for analyzing CGH and gene expression integratedly is not currently available[4,5]. So, we have developed a program called CHESS which detects regions of genomic alteration having susceptibility to a specific phenotype and implements the correlation analysis providing the systematic insight on the genomic alterations and its influences on the level of gene expression.

2 Method and Results

The CHESS has been developed on java application platform and we have used java.math package provided by Apache for implementing statistical engines. CHESS is composed of two main modules: i) detection of phenotype specific CNVs, ii) correlation analysis between CGH and gene expression level. For detection of phenotype specific genomic alteration, CHESS provides three statistical methods, Fisher's exact test, chi-square test and Student's t-test. Simple linear regression,

Spearman rank correlation and Fisher's exact test are available for correlation analysis between CGH and gene expression pattern. Visualization module of CHESS enables users to get the intuitive understanding of analysis results. A frequency plot shows genomic alteration patterns detected in samples, and phenotype-susceptible regions containing statistically significant probes are highlighted in the genome browser. CHESS also tabulates the statistically significant probes and provides gene annotation. Figure 1 shows schematic overview of CHESS.

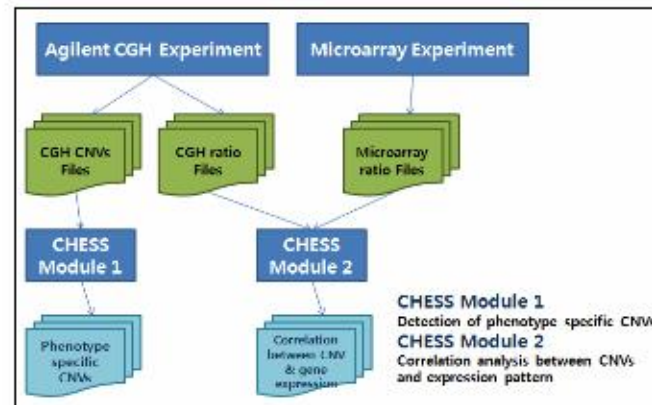


Figure 1. Program Module

3 Discussions

As CHESS provides a comfortable interface to locate the regions of recurrent genomic alteration detected in samples having particular phenotype, it will help shorten the several tedious analysis procedures and save the analysis time for association study between genomic alterations and a particular phenotype. Also, the feature of correlation analysis between CGH ratio and gene expression ratio will provide the clue of up or down regulated genes in the expression level of mRNA induced by genomic alteration patterns. Therefore, the integrative analysis of CGH and gene expression will provide the systematic insight on genomic alterations and their influences in the level of gene expression.

References

- [1] Margolin, A.A. *et al* (2005) CGHAnalyzer: a stand-alone software package for cancer genome analysis using array-based DNA copy number data. *Bioinformatics*, **21**, 3308-3311.
- [2] Lingjærde, O.C. *et al* (2004) CGH-Explorer: a program for analysis of array-CGH data. *Bioinformatics*, **21**, 821-822.
- [3] Shankar, G. *et al* (2006) aCGHViewer: A Generic Visualization Tool For aCGH data. *Cancer Informatics*, **2**, 36-43
- [4] Tsafrir, D. *et al* (2006) Relationship of Gene Expression and Chromosomal Abnormalities in Colorectal Cancer. *Cancer Research*, **66** (4), 2129-2137
- [5] Wieringen, W.N. *et al* (2006) ACE-it : a tool for genome-wide integration of gene dosage and RNA expression data. *Bioinformatics*, **22**, 1919-1920