Molecular expression analysis using DNA microarray and KeyMolnet®

Hajime Fujishima,Shoichi Fumoto,Tsuyoshi Etoh,Masafumi Inomata

Abstract

Background

Recently, neoadjuvant chemotherapy with docetaxel/cisplatin/5-fluorouracil (NAC-DCF) was identified as a novel strong regimen with a high rate of pathological complete response (pCR) in advanced esophageal cancer in Japan. Predicting pCR will contribute to the therapeutic strategy and the prevention of surgical invasion. However, a predictor of pCR after NAC-DCF has not yet been developed. The aim of this study was to identify a novel predictor of pCR in locally advanced esophageal cancer treated with NAC-DCF.

Patients and Methods

A total of 32 patients who received NAC-DCF followed by esophagectomy between June 2013 and March 2016 were enrolled in this study. We divided the patients into the following 2 groups: pCR group (9 cases) and non-pCR group (23 cases), and compared gene expressions between these groups using DNA microarray data and KeyMolnet. Subsequently, a validation study of candidate molecular expression was performed in 7 additional cases.

Results

Seventeen molecules, including transcription factor E2F, T-cell-specific transcription factor, Src (known as “proto-oncogene tyrosine-protein kinase of sarcoma”), interferon regulatory factor 1, thymidylate synthase, cyclin B, cyclin-dependent kinase (CDK) 4, CDK, caspase-1, vitamin D receptor, histone deacetylase, MAPK/ERK kinase, bcl-2-associated X protein, runt-related transcription factor 1, PR domain zinc finger protein 1, platelet-derived growth factor receptor, and interleukin 1, were identified as candidate molecules. The molecules were mainly associated with pathways, such as transcriptional regulation by SMAD, RB/E2F, and STAT. The validation study indicated that 12 of the 17 molecules (71%) matched the trends of molecular expression.

Conclusions

A 17-molecule set that predicts pCR after NAC-DCF for locally advanced esophageal cancer was identified.

Citation: Hajime Fujishima,Shoichi Fumoto,Tsuyoshi Etoh,Masafumi Inomata Molecular expression analysis using DNA microarray and KeyMolnet®. **protocols.io**

dx.doi.org/10.17504/protocols.io.kavcse6

Published: 14 Oct 2017

Protocol

microarray analysis and KeyMolnet analysis

Step 1.

1. **Preparation of RNA and DNA**

Frozen specimens were homogenized, and total RNA was extracted using QIAamp™ DNA Mini Kit (QIAGEN Inc., Valencia, CA) and QIAGEN RNeasy™ mini kit (QIAGEN), according to the manufacture's protocol. Total RNA (200 ng) was reverse transcribed to cDNA using murine leukemia virus reverse transcriptase (Invitrogen Crop., Carlsbad, CA).

2. **Gene expression analysis using microarray analysis**

A human 8 × 60K whole genome oligo DNA microarray chip (SurePrint G3 Human Gene Expression v3 Microarray Kit, G4851C, Agilent Technologies, Santa Clara, CA) was used for global gene expression analysis, according to the manufacturer's protocol. Cyanine (Cy)-labeled cRNA was prepared using T7 linear amplification, according to the Agilent Low RNA Input Fluorescent Linear Amplification Manual (Agilent Technologies). Labeled cRNA was fragmented and hybridized to the same oligonucleotide microarray (Agilent Technologies). The fluorescent intensities were determined with an Agilent DNA Microarray Scanner and analyzed as described using Feature Extraction v.10.7.3.1 (Agilent Technologies). Expression levels were converted into $\log_2$ values and normalized to the median of the entire spot array using GeneSpring™ GX11 (Agilent Technologies). Following normalization, $\log_2$ fold change ($\log_2$FC) in gene expression was calculated using Microsoft Excel® 2016 (Microsoft Corp., Redmond, WA), and the formula was as follows:

Further analysis was performed using KeyMolnet.

3. **Molecular expression analysis using KeyMolnet**

The molecular networks and pathways were analyzed using the KeyMolnet Viewer program version 6.1 (KM Data; www.km-data.jp). KeyMolnet, another commercial knowledge base, contains manually curated contents on 164,000 relationships among human genes and proteins, small molecules, diseases, pathways, and drugs. They include the core contents collected from selected review articles with the highest reliability [1].

KeyMolnet automatically provides corresponding molecules as a node on the networks, by importing the list of Entrez Gene ID and signal intensity data [2, 3]. In this study, gene data, for which expressions were significantly different between the pCR group and non-pCR group, were imported into KeyMolnet. Subsequently, the molecular expressions were calculated and the molecules, which were included in the canonical networks of cancer chemotherapy, were isolated as candidate molecules.

4. **Molecular pathway analysis using KeyMolnet**

To identify the relations of the candidate molecules and canonical pathways, pathway analyses were performed. An algorithm that counts the number of overlapping molecular relations between the extracted network and the canonical pathway allows the identification of the canonical pathway showing the most significant contribution to the extracted network. The significance in the similarity

between both was scored using the following formula:

$(f(x) = {}_{C}C_{x} \square_{T-C}C_{V-x} / {}_{T}C_{V})$

$\text{Score} = -\log_2(\text{Score (p)})$

where O = the number of overlapping molecular relations between the extracted network and the canonical pathway, V = the number of molecular relations located in the extracted network, C = the number of molecular relations located in the canonical pathway, T = the number of total molecular relations (approximately 90,000 sets), and X = the sigma variable that defines incidental agreements [3, 4].

This calculation formula contained the hypergeometric distribution, and the score of more than 20 was considered statistically significant.

1. Satoh J, Asahina N, Kitano S, Kino Y. A Comprehensive Profile of ChIP-Seq-Based Olig2 Target Genes in Motor Neuron Progenitor Cells Suggests the Possible Involvement of Olig2 in the Pathogenesis of Amyotrophic Lateral Sclerosis. J Cent Nerv Syst Dis. 2015;7:1-14.
2. Kuzuhara T, Suganuma M, Kurusu M, Fujiki H. Helicobacter pylori-secreting protein Tipalpha is a potent inducer of chemokine gene expressions in stomach cancer cells. J Cancer Res Clin Oncol. 2007;133(5):287-296.
3. Satoh J, Illes Z, Peterfalvi A, Tabunoki H, Rozsa C, Yamamura T. Aberrant transcriptional regulatory network in T cells of multiple sclerosis. Neurosci Lett. 2007;422(1):30-33.
4. Satoh J, Obayashi S, Misawa T, Sumiyoshi K, Oosumi K, Tabunoki H. Protein microarray analysis identifies human cellular prion protein interactors. Neuropathol Appl Neurobiol. 2009;35(1):16-35.