MOLECULAR ONCOLOGY XXX (2014) 1–20

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/molonc

Transcriptional coexpression network reveals the involvement of varying stem cell features with different dysregulations in different gastric cancer subtypes

Q2 *Kalaivani Kalamohan, Jayaprakash Periasamy, Divya Bhaskar Rao, Georgina D. Barnabas, Pon Srigayatri, Kumaresan Ganesan**

Cancer Genetics Laboratory, Department of Genetics, Centre for Excellence in Genomic Sciences, School of Biological Sciences, Madurai Kamaraj University, Madurai 625021, India

ARTICLE INFO

Article history:
Received 26 December 2013
Received in revised form
22 March 2014
Accepted 14 April 2014
Available online ■

Keywords:
Gastric cancer
Intestinal subtype
Diffuses subtype
Mesenchymal stem cells
Embryonic stem cells
Notch signaling
Polycomb repression
Mesenchymal stem
cell differentiation

ABSTRACT

Despite the advancements in the cancer therapeutics, gastric cancer ranks as the second most common cancers with high global mortality rate. Integrative functional genomic investigation is a powerful approach to understand the major dysregulations and to identify the potential targets toward the development of targeted therapeutics for various cancers. Intestinal and diffuse type gastric tumors remain the major subtypes and the molecular determinants and drivers of these distinct subtypes remain unidentified. In this investigation, by exploring the network of gene coexpression association in gastric tumors, mRNA expressions of 20,318 genes across 200 gastric tumors were categorized into 21 modules. The genes and the hub genes of the modules show gastric cancer subtype specific expression. The expression patterns of the modules were correlated with intestinal and diffuse subtypes as well as with the differentiation status of gastric tumors. Among these, G1 module has been identified as a major driving force of diffuse type gastric tumors with the features of (i) enriched mesenchymal, mesenchymal stem cell like, and mesenchymal derived multiple lineages, (ii) elevated *OCT1* mediated transcription, (iii) involvement of Notch activation, and (iv) reduced polycomb mediated epigenetic repression. G13 module has been identified as key factor in intestinal type gastric tumors and found to have the characteristic features of (i) involvement of embryonic stem cell like properties, (ii) Wnt, MYC and E2F mediated transcription programs, and (iii) involvement of polycomb mediated repression. Thus the differential transcription programs, differential epigenetic regulation and varying stem cell features involved in two major subtypes of gastric cancer were delineated by exploring the gene coexpression network. The identified subtype specific dysregulations could be optimally employed in developing subtype specific therapeutic targeting strategies for gastric cancer.

* *Corresponding author*. Tel.: +91 452 2456224.
E-mail address: kumar@oncocellomics.org (K. Ganesan).

1574-7891/$ – see front matter © 2014 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.molonc.2014.04.005

Please cite this article in press as: Kalamohan, K., et al., Transcriptional coexpression network reveals the involvement of varying stem cell features with different dysregulations in different gastric cancer subtypes, Molecular Oncology (2014), http://dx.doi.org/10.1016/j.molonc.2014.04.005

1. Introduction

Despite the advancements in the cancer research and developments in cancer therapeutics, gastric cancer remains one of the cancers with highest annual mortality rate (Nagini, 2012). Potential therapeutic targets and targeted therapeutics still remain to be identified and developed to tackle different types of gastric cancers. Heterogeneities and complexities remain the major limiting factors in identifying the realistic therapeutic targets (Sehn et al., 2003). Stratification of gastric cancer patients based on the dysregulations and identifying the therapeutic targets for the different subgroups of patients are the current needs in the development of next generation diagnostics and therapeutics. Genome-wide profiling of cancers and integrative functional genomics approaches are found useful in delineating the complexities and heterogeneities in various cancers and molecular stratification of cancer patients (TCGA, 2012; Schroeder et al., 2013).

With the advent of genomic investigations, several advancements have been made in understanding the biology of gastric cancer. Identification of key cancer genes, molecular stratification of tumors, and identification of predominant pathways involved in gastric cancer are the major outcomes (Cho et al., 2011; Ooi et al., 2009; Zang et al., 2011). Interactions among the genes of a single cell type and cross talks among various types of cells in the tumor microenvironment contribute tremendously in cancers and needs to be investigated (Hanahan and Weinberg, 2011). However, these interactions could not be investigated by conventional genomics which often merely involves screening the genes differentially altered in tumors. Integrative and multi-dimensional analysis of microarray data, particularly the transcriptomics data is capable of yielding system level information (Barabasi and Oltvai, 2004; Stuart et al., 2003).

Gene coexpression network based gene pattern analysis from the transcriptomics data by applying the weighted gene coexpression network analysis (WGCNA) organizes the whole genome expression data into functional clusters of coexpressed genes. This is useful to investigate the gene expression profile in functional contexts and to infer the expression pattern of gene sets and their regulations (Horvath and Dong, 2008). Earlier studies involving the construction of coexpression network in gastric cancer had identified coexpressed modular genes and pathways. *PLAG2A*, a prognostic marker, its coexpression with *EPHB2* receptor, association with Wnt/β-catenin pathway and its β-catenin mediated regulation were established from the network of gastric cancer transcriptome (Aggarwal et al., 2006; Ganesan et al., 2008). Another recent network from gastric tumors has revealed the enrichment of the stromal cells in diffuse gastric tumors (Wu et al., 2013). Apart from the mass of proliferating cancer cells, tumors are composed of multiple distinct cell types and the aggressiveness of the cancer is influenced by heterotypic interactions among these cells; in particular, the stromal cells and stem cells contribute to the development and progression of cancers upon differentiation and were inferred from mRNA network (Ben-Porath et al., 2008; Wu et al., 2013; Zhao et al., 2010).

Though multiple networks have been constructed in gastric and many other cancer types, each of these networks have their unique potential in identifying novel system level information in understanding the biology of cancers. In this study, weighted gene coexpression based network analysis of the global transcriptome of gastric tumors was performed to infer the global gene interactions and thus the functional processes playing crucial role in gastric carcinogenesis. It was aimed to connect the gene modules with clinical traits and to understand the gene interactions involved in specific clinical phenotypes. From the coexpression pattern of genes, the major molecular cellular factors involved in two different major subtypes of gastric cancer were identified. Involvements of heterogenous categories of stem cells, varying transcription programs, and different epigenetic dysregulations have been identified as hallmarks of gastric cancer subtypes.

2. Materials and methods

2.1. Microarray data preprocessing

Gene expression profile of gastric tumors was collected from the microarray database Gene expression omnibus (GEO). Since the aim of the study is to obtain the gene network, where each node represents the gene, the MAS 5.0 normalized mRNA profile data was matched with the gene symbol and gene description provided in the corresponding platform file. Gene duplicates were removed by considering the average expression value of multiple probes of genes. The processed profile data was used for the network construction.

2.2. Coexpression network construction

Gastric cancer mRNA profile datasets were collected from Gene expression omnibus (GSE15459, GSE22377) (Forster et al., 2011; Ooi et al., 2009). The first network was constructed from the GSE15459 dataset having the mRNA profile of 200 gastric tumor samples. A weighted coexpression network of the selected 20,318 genes from the gastric cancer transcriptome was constructed by applying the algorithms of WGCNA (Langfelder and Horvath, 2008; Zhang and Horvath, 2005). The correlations among gene expressions were measured based on the Pearson correlation coefficients of all pairs of the genes. The soft thresholding function of WGCNA was used to derive the continuous value of the gene coexpression from Pearson correlation values. The correlation coefficients were further converted into adjacency matrix, using the power adjacency function. The details and scripts used for the construction of coexpression network are provided in Supplemental Method 1.

2.3. Module detection

The adjacency value of each gene along with its degree of shared correlation within the weighted network was transformed into topological overlap, based on which the dissimilarity measure of the topological overlap matrix of the highly coexpressed genes were calculated and the genes

Please cite this article in press as: Kalamohan, K., et al., Transcriptional coexpression network reveals the involvement of varying stem cell features with different dysregulations in different gastric cancer subtypes, Molecular Oncology (2014), http://dx.doi.org/10.1016/j.molonc.2014.04.005