Microbial Pathogenesis 105 (2017) 100–105

Contents lists available at ScienceDirect

Microbial Pathogenesis

journal homepage: www.elsevier.com/locate/micpath

Immune response- and viral control-related pathways in the progression of chronic hepatitis B

Zhi-Qiang Zou [a, 1], Shuai Zhang [b, 1], Qing Lin [b], Ren-Liang Qu [b], Yan-Fang Li [c], Fu-Hua Zhang [b], Ai-Ling Xu [b, *]

[a] Yantai City Hospital for Infectious Diseases, Yantai 264001, China
[b] Clinical Laboratory, Yantai City Hospital for Infectious Diseases, Yantai 264001, China
[c] Department of Hepatobiliary Internal Medicine, Yantai City Hospital for Infectious Diseases, Yantai 264001, China

ARTICLE INFO

Article history:
Received 9 January 2017
Received in revised form
8 February 2017
Accepted 8 February 2017
Available online 9 February 2017

Keywords:
Chronic hepatitis B
Differential co-expression network
Sub-network
Host response

ABSTRACT

Background: Chronic hepatitis B (CHB) is a complicated and dynamic course, and is associated with advanced liver disease. Host immune response against viral infection plays a pivotal role in the progression of CHB. However, it is still uncharted that how the hepatic transcriptomes in patients with CHB are correlated with the clinical phases.

Objective: This study aimed to identify the specific sub-networks across various phases of CHB and infer potential pathways for phenotypic outcome prediction.

Methods: In this study, we performed the pairwise comparisons of the hepatic transcriptomes of CHB patients under different phases, and constructed the differential co-expression networks (DCNs). We firstly identified the critical genes from each DCN according to the adjacency matrix of the network. Then, the specific sub-networks were digged by iteratively affiliating genes that can increase the classification accuracy, using a snow-ball sampling strategy. Permutation test was implemented to determine the statistical significance of these sub-networks. Finally, each sub-network was given a most significant functional pathway.

Results: We constructed 3 DCNs by pairwise comparing the hepatic transcriptomes among three CHB phases, and systemically tracked 1, 1 and 2 specific sub-networks and pathways, respectively. Relative to immune tolerant phase, TGF-beta receptor signaling in EMT (epithelial to mesenchymal transition) pathway was significantly changed in the immune clearance phase, and nuclear receptor transcription pathway and adenylate cyclase activating pathway were altered in inactive carrier state. The host genes related to DNA strand elongation showed significant difference between the immune clearance phase and inactive carrier state.

Conclusions: By pairwise comparing the hepatic transcriptomes of CHB patients under a network view, several immune- and viral control-related pathways were identified in this study. These results might serve as a foundation for characterizing the host transcriptomes responded to CHB infection, and hold clues for the development of potential targets for disease control.

1. Introduction

Hepatitis B virus (HBV) infection is a challenging global health concern, which can cause a potentially life-threatening liver failing. Stupendously, about one third of the world population has been infected at one point in their lives [1]. World Health Organization estimates that approximate 240 million persons worldwide are chronic carriers of HBV, i.e. chronic hepatitis B (CHB), and an estimated 650,000 people will die annually due to CHB-associated liver disease, such as cirrhosis and liver cancer [2,3]. Of those CHB individuals worldwide, one third reside in China [4,5]. Current therapies for CHB are limited to interferon-α and various nucleosides, that could reduce viremia and improve long term outcome [6]. Despite advances in treatment, limited effects have displayed currently, because it is a silent disease until it is complicated by

* Corresponding author. Clinical Laboratory, Yantai City Hospital for Infectious Diseases, No. 62 Huanshan Road, Zhifu District, Yantai 264001, China.
E-mail address: 15564536548@163.com (A.-L. Xu).
[1] Equal contributors and co-first authors.

http://dx.doi.org/10.1016/j.micpath.2017.02.015
0882-4010/

advanced liver disease. Therefore, an improved understanding of the natural history of CHB will contribute to the guidance of management and optimal timing for initiating antiviral treatment of CHB, avoiding progression to advanced phases.

CHB is a complicated and dynamic course. The natural course of CHB has three phases: an immune tolerance phase with minimal liver disease, immune clearance phase with active liver inflammation, an inactive carrier state with minimal inflammation and fibrosis on liver biopsy. Normally, patients with CHB can progress from one phase to the next, while may revert backward [7]. Accumulating evidence recognizes that the confrontation between of host immune responses and viral replication plays a pivotal role in the development of HBV-related liver disease [7,8]. Understanding the natural history of CHB will contribute to the guidance of optimal timing for initiating antiviral therapy. However, it is still uncharted that how the intrahepatic transcriptomes in patients with CHB are correlated with the clinical phases. A better understanding of the biological processes across various phases of CHB could facilitate the efficient management and therapies.

It has been largely accepted that genetic factors play a vital role in the pathogenesis of CHB. Transcriptome profiling analyses indicate that the biological processes of diseases are associated with the dys-regulation of specific groups of genes and their interconnections [9,10]. Network-based approach shows great potential for uncovering disease-related genes and pathways [11]. Meanwhile, sub-network markers could not only cover critical genes that are functionally related to diseases, but also predict the clinical outcomes accurately [12,13].

In the present study, we attempted to identify the specific sub-networks across various phases of CHB by conducting network analysis on the intrahepatic transcriptomes of CHB patients responded to viral infection. The results might contribute to a better understanding of underlying pathological mechanism in the process of CHB, and infer potential signatures for phenotypic outcome prediction.

2. Materials and methods

2.1. Intrahepatic transcriptome data

The intrahepatic transcriptome data were recruited from ArrayExpress database (http://www.ebi.ac.uk/arrayexpress/), under the accessing number of E-GEOD-65359. The transcriptome data were performed on Affymetrix GeneChip Human Genome U133 Plus 2.0 array. In this dataset, there are a total of 83 liver biopsy samples from patients with CHB, including 22 immune tolerance, 50 immune clearance, and 11 inactive carrier state. In this study, we performed the pairwise comparisons of the intrahepatic transcriptomes of three phases. Prior to primary analysis, standard data preprocessing was performed to control the quality of the transcriptome data, including background correction [14], normalization [15], and probe correction [16]. Then, the data on probe levels were summarized by medianpolish [14]. Each probe was mapped to one gene symbol, and the probe was discarded if it could not match any genes. Finally, a total of 13,326 genes and their corresponding expression information were obtained for subsequent analysis.

2.2. Differential co-expression networks

Differential co-expression analysis is a powerful tool to explore diagnostic gene signatures and biological processes of complex diseases. Differential co-expression analysis focuses on the potential interactions among genes and the gene co-expression patterns across different conditions, contributing to the understanding of altered gene regulatory mechanisms [17]. To identify the sub-network markers for predicting the clinical outcomes, it is indispensable to construct the differential co-expression network (DCN). To achieve this, the global human protein-protein interactions (PPIs) were firstly downloaded from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (http://string-db.org/) that is a database of known and predicted PPIs from various sources of evidence. There are many self-interacting proteins with self-link loops that might affect the conservation and topology of PPI network. In this study, we ignored the self-loops and duplicated interactions, and gained a total of 787,896 interactions (covering 16,730 genes). Here we extracted the common genes between transcriptome data and PPI data, and remained 7953 genes and 48,778 interactions for subsequent analysis.

In this study, we performed the pairwise comparisons of the intrahepatic transcriptomes of three phases and could constructed three DCNs, named as DCN1 (immune clearance vs immune tolerance), DCN2 (inactive carrier state vs immune clearance), and DCN3 (inactive carrier state vs immune tolerance). The DCNs were constructed based on co-expression analysis and differential expression analysis successively. Firstly, the Pearson correlation coefficients (PCC) [18] was employed to characterize the co-expression levels of gene pairs under various disease conditions, and only PPIs with $|PCC| > 0.8$ were remained to build the co-expression network. Then, we implemented a two-tailed *t*-test to determine the differential expression between two conditions, with Benjamini-Hochberg false discovery rate (FDR) adjustment [19]. Finally, a weight value was assigned to each interaction by a Bioconductor package edgeR [20], to weigh the degree of the co-expression and differential expression. Under this case, each edge in DCNs was assigned a weight value, and the interaction with high co-expression and high differential expression was given to a high weight value.

2.3. Prediction of critical genes

After constructing DCNs, we tried to determine the critical genes from each DCN by degree centrality analysis of network. Based on the degree value of each gene, a z-score [21] was bestowed on each gene according to the adjacency matrix of the network. The significances of the genes in each DCN were measured according to their z-scores. All genes were ranked in descending order based on their z-score values, and the top 1% genes with the largest z-scores were defined as critical genes of each DCN.

2.4. Identification of diagnostic sub-networks

In this section, we attempted to identify diagnostic sub-networks with maximum classification accuracy from each DCN, respectively, using a snow-ball sampling strategy [22]. Starting from each critical gene, candidate diagnostic sub-network search was conducted by critical gene expansion. In this process, it iteratively affiliated genes whose accession led to the maximum increase in the prediction accuracy model until the prediction accuracy dropped. To assess the prediction accuracy, the area under the receiver operating characteristic curve (AUC) was employed as the metric. In this study, the sub-networks with AUC >0.9 as well as gene size ≥ 5 were considered as candidate sub-networks.

Subsequently, permutation test was implemented to determine the statistical significance of candidate sub-networks. We firstly constructed randomized networks based on the null score distribution, and searched sub-networks from the randomized networks using the method mentioned above. Each procedure was completely randomized for 1000 permutations, and each sub-network was assigned an AUC score in each one permutation. The