Matrix Biology 32 (2013) 393–398

Contents lists available at ScienceDirect

Matrix Biology

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

Brief report

Extracellular matrix transcriptome dynamics in hepatocellular carcinoma

Michael B. Duncan *

Section of Gastroenterology/Hepatology, Department of Medicine, Medical College of Georgia, Georgia Regents University, Augusta, GA 30912, United States
Department of Biochemistry and Molecular Biology, Medical College of Georgia, Georgia Regents University, Augusta, GA 30912, United States
Georgia Regents University Cancer Center, Georgia Regents University, Augusta, GA 30912, United States

ARTICLE INFO

Article history:
Received 3 April 2013
Received in revised form 17 May 2013
Accepted 21 May 2013

Keywords:
Cancer
RNA-Seq
Collagen
Proteoglycan
Glycoprotein

ABSTRACT

The extracellular matrix undergoes extensive remodeling during hepatocellular carcinoma and functions as a critical component of the tumor microenvironment by providing a substratum for cell adhesion and serving as a reservoir for a variety of cytokines and growth factors. Despite the clinical correlation between ECM deposition and hepatocellular carcinoma progression, it remains unclear how global extracellular matrix gene expression is altered in hepatocellular carcinoma and the molecular pathways that govern this change. Herein, a comprehensive analysis of the extracellular matrix transcriptome using an RNA-sequencing dataset provided by The Cancer Genome Atlas consortium was conducted and indicates substantial differential gene expression of key extracellular matrix collagens, glycoproteins, and proteoglycans in hepatocellular carcinoma. This analysis also reveals alternative expression of extracellular matrix gene transcript variants that could impact biological activity and serves as a framework for exploring the dynamic nature of the extracellular matrix transcriptome in cancer and identifying candidate genes for future exploration.

1. Introduction

It is estimated that 80–90% of all cases of hepatocellular carcinoma are preceded by cirrhosis (Seitz and Stickel, 2006). In cirrhosis, reduced liver function is due in part to the excessive deposition of extracellular matrix (ECM) and activation of various stromal cell types. These changes in the ECM also provide a permissive environment for tumor development through enhanced integrin signaling, paracrine cross-talk between tumor and stromal cells, increased tissue stiffness, and altered growth factor sequestration by ECM (Zhang and Friedman, 2012). This suggests that HCC progression and clinical outcomes are tightly linked to the tissue remodeling and enhanced matrix deposition within the stromal tissue environment that typifies cirrhotic liver pathology. From a clinical perspective, curative therapy for HCC when moderate to extensive cirrhosis has occurred is often not effective (Ercolani et al., 2003). These observations provide strong evidence that the extracellular matrix in the hepatic microenvironment plays a critical role in HCC progression and can hamper therapeutic outcomes.

The ECM is a fundamental component of the tumor stroma, promoting cancer hallmark capabilities to tumor cells via enhanced cell adhesion and serving as a reservoir for growth factors and cytokines (Hanahan and Weinberg, 2011; Lu et al., 2012). Due to its distinct molecular composition, the liver ECM is structurally unique from most tissue ECMs and lacks the typical electron dense nature of most basement membrane structures. Quantitative and qualitative studies convincingly show that as liver tissue changes during cirrhosis, so does the composition of the ECM (Gressner and Weiskirchen, 2006; Zeisberg et al., 2006). However questions remain regarding the origins of these changes and their impact on cancer progression. In particular how the ECM is altered in HCC and the molecular pathways that govern this change are largely unknown. Proteomic approaches have been developed to begin determining the types of ECM and ECM related molecules expressed in the tumor microenvironment of various cancers (Hynes and Naba, 2012; Naba et al., 2012). The genetic programs and resulting expression that give rise to these unique ECM profiles in these and other cancer types remain to be fully elucidated.

Next generation sequencing represents a significant advance in our ability to study genomic alterations and global gene expression profiles in cancer (Shendure and Ji, 2008; Han, 2012). In particular, RNA sequencing technology provides sensitive detection of gene expression, analysis of non-coding RNA, and the discovery of novel transcripts in cancer (Mardis and Wilson, 2009; Wang et al., 2009). Recently, The Cancer Genome Atlas (TCGA, http://cancergenome.nih.gov/) has initiated an effort to provide publicly accessible data on several cancers including HCC. In this study the TCGA dataset was used to develop a preliminary assessment of changes in the extracellular matrix transcriptome in hepatocellular carcinoma. These

* Medical College of Georgia, Georgia Regents University, Augusta, GA 30912, United States. Tel.: +1 706 721 5484.
E-mail address: mduncan@gru.edu.

0945-053X/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.matbio.2013.05.003

results provide new insight into the ECM gene expression pattern in the liver and the substantial changes it undergoes during HCC.

2. Results

2.1. ECM expression in normal versus HCC

The TCGA provides gene expression datasets (RNA sequencing) for a variety of cancers including HCC (LIHC dataset). For this study, normalized RNA counts derived from the RSEM (RNA-Seq expression estimation by Expectation–Maximization) algorithm were acquired from the TCGA data portal and analyzed. The RSEM algorithm is a statistical model that estimates RNA expression levels from RNA sequencing counts (Li and Dewey, 2011). Recently a comprehensive list of ECM and ECM related molecules has been reported (Hynes and Naba, 2012; Naba et al., 2012). These studies have classified ECM genes into "core" and "associated" ECM subtypes, in which the "core" sub-type refers to proteins that are directly incorporated into the ECM structure, while the "associated" sub-type refers to proteins that interact with or alter the ECM. This list of "core" ECM genes was utilized to conduct an analysis of ECM gene expression in human HCC. ECM gene expression was assessed based on the RSEM counts. A broad distribution of ECM gene expression was observed as indicated by RSEM normalized counts in both normal liver (n = 20) and HCC (n = 34) samples (Fig. 1A). More ECM genes were expressed at greater than 50 counts in HCC as compared to normal, suggesting a general increase in ECM gene expression. A scatter plot comparing gene expression between normal and HCC reveals a broad spectrum of expression among collagens, glycoproteins, and proteoglycans of the ECM (Fig. 1B). Based on the scatter plot nearly a third of the analyzed ECM genes were potentially alternatively expressed in HCC.

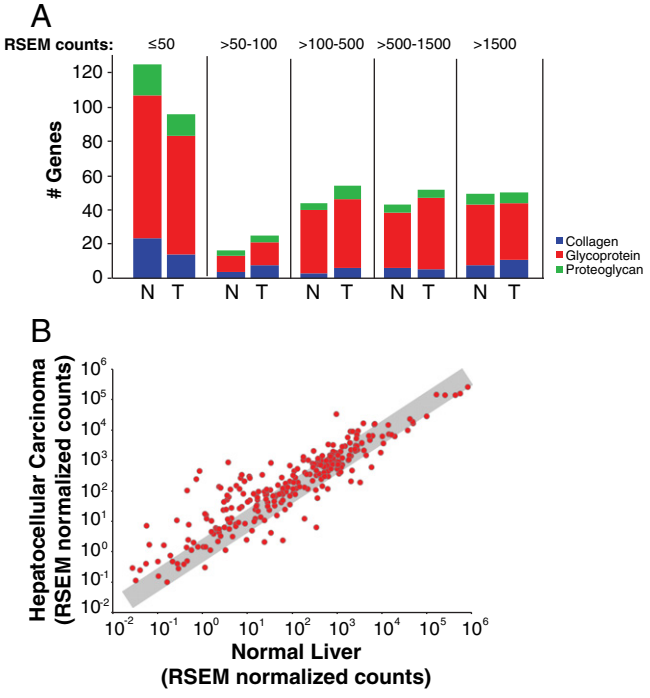

Fig. 1. ECM gene expression in normal liver (N) versus hepatocellular carcinoma (T) tissue. (A) RSEM normalized count distribution in normal liver (n = 20) and hepatocellular carcinoma (n = 34). (B) Scatter plot of RSEM normalized gene expression in normal liver and hepatocellular carcinoma tissues. Normalized values were averaged and plotted on a log scale. The gray shaded area indicates less than two fold difference in RSEM normalized value between normal and HCC tissue.

2.2. Differential expression of ECM in HCC

In order to establish the role of the ECM in HCC, it is essential to determine what ECM genes are differentially expressed in the tumor setting. To determine specific genes that were significantly differentially expressed, a volcano plot of the statistical difference between normal and HCC samples (p-value of Student's *t*-test) versus fold change in gene expression of normal versus HCC was analyzed (Fig. 2A). Based on the volcano plot (see Experimental procedures for detail), a diverse set of ECM genes was differentially expressed (both over and under expressed) in HCC. Fig. 2B depicts the number of collagens, glycoproteins and proteoglycans that are up or down regulated or did not change significantly during HCC. Criteria of a two-fold difference in gene expression and a $p \geq 0.05$ were set in order to establish differential expression. An assessment of the three major ECM gene families indicate that 40.91%, 18.97%, and 25.00% of collagens, glycoproteins, and proteoglycans are up-regulated respectively and 6.82%, 15.38%, and 5.56% of collagens, glycoproteins, and proteoglycans were down regulated in HCC.

2.3. Highly expressed ECM genes in HCC

In order to establish potential roles of ECM genes that are differentially expressed in HCC, it is important to determine if they are expressed at significantly high levels. ECM genes were plotted based on their RSEM normalized counts. A threshold of greater than 1500 counts was established in order to focus on those genes that were expressed to an appreciable extent in the liver. It is important to note that a number of differentially expressed ECM genes fell below this established threshold and a complete list of all ECM genes analyzed in the dataset is provided in Supplemental Table 1. We observed that several collagens including *COL1A1*, *COL1A2*, *COL4A1*, *COL4A2*, *COL5A1*, *COL5A2*, and *COL6A3* met this threshold and were up regulated in HCC (Fig. 3A). No collagens meeting the expression threshold were down regulated. The proteoglycans *AGRN*, *HSPG2*, and *VCAN* were up regulated in HCC while *DCN* and *PRG4* proteoglycans met the expression threshold but were down regulated in HCC (Fig. 3B). ECM glycoproteins that met the expression threshold including *CYR6*, *ECM1*, *FGA*, *FGB*, *FGG*, *FGL1*, *IGFALS*, *IGFBP1*, *IGFBP3*, *IGFBP4*, *LRG1*, *MFAP3L*, and *THBS1* were down regulated, while *SMOC1*, *SPARCL1*, *SPARC* and *VWF* were up regulated in HCC (Fig. 3C).

2.4. Isoform variants of differentially and highly expressed ECM genes

Isoform variants of ECM genes influence their localization and function. RNA sequencing allows for determining the relative contribution of these isoforms. We examined how isoform levels change during differential expression in HCC. Of the ECM genes that were down regulated in HCC, *FGL1* and *MFAP3L* demonstrated switches in the relative abundance of their predominant isoforms (Fig. 4A). Of the up-regulated genes that met the expression criteria *VCAN*, *HSPG2*, *COL6A3*, and *SPARCL1* showed a significant change in isoform abundance in HCC (Fig. 4B).

3. Discussion

3.1. Extracellular matrix and hepatocellular carcinoma

It is well appreciated that the ECM plays critical roles in developmental, homeostatic, and tumorigenic processes in liver biology (Bedossa and Paradis, 2003; Kalluri, 2003; Hynes, 2009; Yurchenco, 2011). However, a systematic understanding of the composition of the HCC ECM and how it is regulated has not yet been fully elaborated. Understanding the molecular constituents of the ECM as well as the gene expression programs that regulate their expression will