Discovering a critical transition state from nonalcoholic hepatosteatosis to nonalcoholic steatohepatitis by lipidomics and dynamical network biomarkers

Rina Sa1,2,†, Wanwei Zhang2,3,†, Jing Ge2,3, Xinben Wei1,2, Yuhua Zhou1, David R. Landzberg4, Zhenzhen Wang1, Xianlin Han5, Luonan Chen2,3,6,* and Huiyong Yin1,2,6,7,*

1 Key Laboratory of Food Safety Research, Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences (SIBS), Chinese Academy of Sciences (CAS), Shanghai 200031, China
2 University of the Chinese Academy of Sciences, CAS, Beijing 100049, China
3 Key Laboratory of Systems Biology, Innovation Center for Cell Signaling Network, Institute of Biochemistry and Cell Biology, SIBS, CAS, Shanghai 200031, China
4 Vanderbilt School of Medicine, Nashville, TN 37235, USA
5 Diabetes and Obesity Research Center, Sanford – Burnham Medical Research Institute, Orlando, FL 32827, USA
6 School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China
† These authors contributed equally to this work.
* Correspondence to: Huiyong Yin, E-mail: hyyin@sibs.ac.cn; Luonan Chen, E-mail: lnchen@sibs.ac.cn

Nonalcoholic fatty liver disease (NAFLD) is a major risk factor for type 2 diabetes and metabolic syndrome. However, accurately differentiating nonalcoholic steatohepatitis (NASH) from hepatosteatosis remains a clinical challenge. We identified a critical transition stage (termed pre-NASH) during the progression from hepatosteatosis to NASH in a mouse model of high fat-induced NAFLD, using lipidomics and a mathematical model termed dynamic network biomarkers (DNB). Different from the conventional biomarker approach based on the abundance of molecular expressions, the DNB model exploits collective fluctuations and correlations of different metabolites at a network level. We found that the correlations between the blood and liver lipid species drastically decreased after the transition from steatosis to NASH, which may account for the current difficulty in differentiating NASH from steatosis based on blood lipids. Furthermore, most DNB members in the blood circulation, especially for triacylglycerol (TAG), are also identified in the liver during the disease progression, suggesting a potential clinical application of DNB to diagnose NASH based on blood lipids. We further identified metabolic pathways responsible for this transition. Our study suggests that the transition from steatosis to NASH is not smooth and the existence of pre-NASH may be partially responsible for the current clinical limitations to diagnose NASH. If validated in humans, our study will open a new avenue to reliably diagnose pre-NASH and achieve early intervention of NAFLD.

Keywords: nonalcoholic fatty liver disease (NAFLD), mass spectrometry lipidomics, systems biology, pre-NASH, dynamical network biomarkers

Article

Prevalence of NAFLD is estimated to be 30% of the general population in western countries (Browning et al., 2004; Bedogni et al., 2005) and ~27% in some Chinese cities (Fan, 2013). It is estimated that NAFLD can soon become the next global epidemic if the current trends are not reversed (Morgan, 2014). Although the high prevalence of NAFLD has been well recognized, diagnosis and management of NAFLD remain a clinical challenge. It is generally accepted that simple hepatic steatosis has a benign clinical consequence while NASH has a much higher risk to progress to cirrhosis and even hepatocellular carcinoma (Adams et al., 2005). Therefore, it is of utmost importance to clinically differentiate NASH from steatosis.

Introduction

Nonalcoholic fatty liver disease (NAFLD) covers an entire pathological spectrum of liver abnormalities including simple steatosis, nonalcoholic steatohepatitis (NASH), advanced fibrosis, and cirrhosis in the absence of significant alcohol consumption or other known liver diseases (Angulo, 2002; Ray, 2013). NAFLD is now considered as a hepatic manifestation of metabolic syndrome including insulin resistance, type 2 diabetes mellitus, obesity, and cardiovascular diseases.

Prevalence of NAFLD is estimated to be 30% of the general population in western countries (Browning et al., 2004; Bedogni et al., 2005) and ~27% in some Chinese cities (Fan, 2013). It is estimated that NAFLD can soon become the next global epidemic if the current trends are not reversed (Morgan, 2014). Although the high prevalence of NAFLD has been well recognized, diagnosis and management of NAFLD remain a clinical challenge. It is generally accepted that simple hepatic steatosis has a benign clinical consequence while NASH has a much higher risk to progress to cirrhosis and even hepatocellular carcinoma (Adams et al., 2005). Therefore, it is of utmost importance to clinically differentiate NASH from steatosis.
Imaging tools such as liver ultrasonography, computed tomography (CT), and magnetic resonance imaging (MRI) have been used non-invasively to diagnose the presence of hepatic steatosis with a relative high degree of accuracy but limited sensitivity. However, none of these techniques can readily distinguish NASH from simple steatosis. Liver biopsy remains the gold standard for diagnosis and staging of NAFLD (Nalbantoglu and Brunt, 2014). Due to the invasive nature and high cost associated with liver biopsy, extensive efforts have been made to improve the diagnosis and staging of NAFLD with non-invasive methods. Recent focus has shifted to utilize high-throughput data including genomics, proteomics, and metabolomics to search for better biomarkers in serum/plasma (Younossi et al., 2005). In particular, lipidomics, a systematic analysis of all lipids in a given biological system, has attracted increasing interest in this endeavor (Puri et al., 2007, 2009; Lamaziere et al., 2012). High expectations have been given to lipid biomarkers in serum or plasma, which can intimately track the disease progression in the liver. However, limited success has been achieved thus far in this field due to the complex and heterogeneous nature of NAFLD (Wree et al., 2013; Gorden et al., 2015). Moreover, it remains unknown whether the lipid change in the circulation correlates with that in the liver during NAFLD progression.

Current paradigms of disease biomarker discovery are based on the capability to detect distinct production of a molecule or a group of molecules between normal and disease states. Due to the multiple stages of NAFLD and the heterogeneity of its progression, unambiguous differentiation of simple steatosis from NASH represents a tremendous challenge. The Nonalcoholic Steatohepatitis Clinical Research Network in the United States proposed a histological scoring system, termed NAFLD activity score (NAS), to evaluate semi-quantitatively the degrees of steatosis, lobular inflammation, hepatocellular ballooning, and fibrosis in liver tissues. According to the initial proposal, a ‘not NASH’ is diagnosed when NAS ≤ 3 while NASH is diagnosed when NAS ≥ 5 (Kleiner et al., 2005). This scoring system has been routinely employed in clinical diagnosis and in multiple cohorts to define endpoints in clinical trials (Kleiner et al., 2005; Neuschwander-Tetri et al., 2015). However, subsequent studies suggest that this system still cannot definitively diagnose NASH or the absence of NASH based on the threshold value of NAS, especially when the NAS scores lie in the borderline (Brunt et al., 2011). The underlying molecular mechanisms for this difficulty remain poorly understood. It is conceivable that a detailed profile study of lipid signature in blood or liver biopsy using lipidomics may provide better insights into the biochemical bases to differentiate NAS from simple steatosis.

Rapid progress at the front of omic technology enables discovery of novel disease biomarkers reflecting changes of an entire metabolic network (termed network biomarkers) (Zeng et al., 2014; Zhang et al., 2015). Recently, a novel model-free method based on nonlinear dynamic theory, termed dynamical network biomarkers (DNB), was developed to characterize critical transition (or early-warning signals) during the progression of complex diseases (Chen et al., 2012; Liu et al., 2012, 2013, 2014; Zeng et al., 2014) in a completely different way from traditional biomarkers. DNB is a group of molecules with strong collective fluctuations, appearing only at the ‘tipping point’ of a homeostatic system. Specifically, DNB is composed of a group of metabolites whose standard deviations (SDs) and Pearson’s correlation coefficients (PCCs) are both drastically increased at a critical stage just before the transition during the disease progression. A quantitative index, termed criterion index (CI), for the DNB can be derived using mathematic operations (vide infra), and intuitively it measures the differential correlations and deviations of molecular expressions rather than the differential expressions adopted in the traditional methods. Theoretically, the members of the DNB are more likely to be ‘drivers’ of the diseases and thus interventions targeting these members at this critical stage may alter the disease progression (Liu et al., 2012; Li et al., 2014).

In this work, combining mass spectrometry (MS)-based shotgun lipidomics with DNB approach, we identified a critical stage during the progression of NAFLD from simple steatosis to NASH. We tentatively term this stage pre-NASH, which is characterized by (i) a sudden loss of lipid correlations between blood and liver, a hallmark ‘tipping point’ during disease progression; (ii) the maximal quantitative index for DNB lipid species in liver and blood, especially for triacylglycerol (TAG). Interestingly, a majority of DNB members of TAGs in the blood overlapped with those in the liver during the progression of NAFLD, a stark contrast to the poor correlations of the same lipid species between the blood and liver tissue based on abundance. Further metabolic pathway analysis reveals that fluctuation of gene expression related to hepatic pathways of TAG synthesis and degradation precedes the transition from simple steatosis to NASH. In particular, TAG synthesis enzyme, diacylglycerol acyltransferase (DGAT), and hydrolysis enzymes, lipoprotein lipase (LPL) and adipocyte triacylglycerol lipase (ATGL), appear to play a critical role in this transition. Interestingly, collective fluctuation of this inherent process of TAG metabolism echoes recent findings in genome wide association studies (GWAS) in humans in which a TAG hydrolyase patatin-like phospholipid domain containing protein 3 (PNPLA3) is identified as one of the major genetic modifiers influencing NAFLD progression (Anstee and Day, 2013).

**Results**

*A mouse model of NAFLD with high-fat feeding mimics the progression of nonalcoholic hepatosteatosis to NASH based on NAS*

In a high-fat-induced NAFLD model, mice were fed by a high-fat diet (HFD) containing 60% of calorie consumption from fat for 6–18 weeks (Supplementary Figure S1A). This well-established model has been shown to mimic early stages of human NAFLD from normal liver phenotype to hepatic steatosis and NASH (Kanuri and Bergheim, 2013). After animals were sacrificed, blood and liver tissue were harvested every 2 weeks. Based on the hematoxylin and eosin (H&E) staining of liver tissue, NAS was determined and the mice were grouped by their corresponding NAS. The H&E staining showed that the livers of HFD-fed mice covered NAS from 1–5 while the livers of normal diet fed mice were determined as NAS 0 (Figure 1A and Supplementary Figure S1B). As shown in Figure 1B, a gradual increase of total hepatic TAG was observed and this trend appeared to slow down after NAS was greater than 3, consistent with histological classifications. However, the levels of TAG in plasma did not show NAS-dependent increase. This discrepancy has been commonly observed in animal models as
well as in human subjects, highlighting the difficulty to diagnose NAFLD by assessing blood TAG levels. Consistent with the histological assessment and NAS classification, plasma alanine aminotransferase (ALT) and aspartate aminotransferase (AST) were significantly elevated at NAS 5 (Figure 1C). We also observed other common features associated with metabolic syndrome including increased plasma total cholesterol (TC) and fasted glucose levels. High-fat feeding also resulted in significant weight-gain compared with the normal diet fed group (Supplementary Figure S1C and D). In summary, the mouse model of high-fat feeding mimics the early stages of NAFLD in mice and humans; NAS scoring system can be readily applied to classify disease progression in this rodent model.

Figure 1 Histological and biochemical analyses of mouse samples of high fat-induced NAFLD. (A) Representative pictures of histological classification of mouse liver tissue based on H&E staining with a magnification of 200- and 400-fold, respectively. (B) Levels of TAG in the liver and plasma analyzed by biochemical kits. (C) Levels of plasma ALT and AST assayed by biochemical kits. All values are expressed as mean ± SEM. *P < 0.05, **P < 0.01, ***P < 0.001 when compared each NAS group with NAS 0 (n = 6–6 for each group). TAG, triacylglycerol; TC, total cholesterol; ALT, alanine aminotransferase; AST, aspartate aminotransferase.

Lipidomic approach provides better insights into the lipid changes in mouse liver and blood during the progression of NAFLD

Dyslipidemia is a hallmark of NAFLD and lipids have been predicted to be potential biomarkers for the progression of this disease. MS-based lipidomics has emerged as an ideal tool to investigate the lipid metabolism at a system level during the progression of NAFLD in animal models and humans (Puri et al., 2007, 2009; Gorden et al., 2011). To uncover the early metabolic and lipidomic changes associated with HFD-induced NAFLD, we applied a well-established shotgun lipidomic approach to examine molecular lipid species in mouse plasma and liver. This method can routinely detect and quantify around 400 individual lipid ions in
mouse plasma and liver including phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylinositol (PI), phosphatidylserine (PS), phosphatic acid (PA), phosphatidylglycerol (PG), lysophosphatidylcholine (LPC), lysophosphatidylyethanolamine (LPE), sphingomyelin (SM), ceramide (Cer), and TAG. Consistent with the histological classification and total TAG shown in Figure 1, we observed a significant increase of most TAG ions in liver tissue with NAS 4 and 5 compared with NAS 0-3. However, there was no such trend in plasma TAG species (Figure 2A), consistent with no obvious change in total TAG in plasma (Figure 1B). It is interesting to note that the major phospholipids PC and PE in liver tissues were significantly decreased in NAS 1-5 compared with NAS 0; the opposite trends showed in the plasma (Supplementary Figure S2A and B). We next performed Pearson’s correlation analysis on the same lipid species detected both in liver and plasma. To our surprise, we observed a sudden loss of correlations of TAG ions at NAS 3 (Figure 2B). A similar trend was also seen for all lipid species (Figure 2C) and total fatty acids (Supplementary Figure S3), and this abrupt drop in correlations was highly statistically significant. These preliminary data suggest that progression from steatosis to NASH may not be smooth and there may exist a critical transition stage during this process. Further analysis by hierarchical clustering of differentially produced lipid species with statistical significance between steatosis and NASH in plasma and liver tissue showed that these lipid ions can be well separated according to NAS except for NAS 3. NAS 3 (red color) was not readily clustered in liver tissue (Figure 2D) nor in plasma (Figure 2E). Taking these data together, lipidomic approach provides unparalleled molecular details during the progression of NAFLD in mouse liver and plasma, which can be employed to differentiate simple steatosis (NAS ≤ 3) and NASH (NAS ≥ 5) based on the abundance of lipid ions, especially for TAGs. However, classification becomes ambiguous at the critical stage of the disease at NAS 3, which implies that NAS 3 is a special stage. The sudden loss of correlations at NAS 3 between liver and plasma lipids implicates that as NAFLD progresses, changes in abundance of individual or total lipids in circulation may not reflect those in the liver, especially when NAS approaches 3, a putative critical stage, which is the state just before the transition from steatosis to NASH. We tentatively designate this state as pre-NASH state.

**DNB theory further supports the presence of the critical stage from simple steatosis to NASH**

Identifying the critical stage or pre-NASH state during disease progression is crucial for diagnosis and treatment purpose. However, the critical state is difficult to be detected based on abundances of molecular species because there may be no significant differential expressions (abundances) of molecules between steatosis and pre-NASH states (Figure 3A and B). To address this issue, a DNB model was developed by measuring collective fluctuations of molecules rather than their differential expressions, and this is the first such theoretical work to detect the pre-disease state or critical stage (see Methods). The three criteria for DNB members or molecules are summarized in Figure 3B, i.e. at the critical stage, (i) the expression of these molecules in DNB becomes highly fluctuated (high SDs), (ii) these molecules in DNB are highly correlated as indicated by high PCCs, and (iii) correlations between DNB members and all other non-DNB species disappear (low PCCs). The features of the DNB molecules and the interactions among these molecules can be depicted in a network manner (Figure 3C). The prerequisite for this method is to obtain high-throughput omic data (with at least three samples) that cover the entire molecular network of interest. Traditional methods use the differential expressions of molecules while DNB exploits both the differential correlations and differential deviations among molecules. However, as shown in Figure 3B, there are no significantly differential expressions between steatosis and pre-NASH states, but there are significantly differential correlations and deviations between those two states based on DNB theory, which is why DNB can detect the pre-NASH state. Our preliminary results suggest that there is a potential critical stage from simple steatosis to NASH (NAS 3). We next applied the DNB method to confirm the existence of this state. Specifically, based on the lipidomic profiles obtained by MS of mouse plasma and liver tissue, we identified lipid class specific and total DNB members in plasma and liver tissue corresponding to the critical state of NAFLD progression. Since TAGs represent the most diverse class of lipids and were shown to increase most significantly in plasma and liver tissue during NAFLD progression, we first screened DNB members of TAGs in mouse plasma and liver respectively. Thirty-nine and sixty-four TAG species were identified as DNB members in plasma and liver respectively (shown in bold in Table 1). Interestingly, TAGs in plasma and liver underwent the same critical transition just after NAS 3 as indicated by a maximal CI value, suggesting that this critical stage is consistent across plasma and liver tissue (Figure 4A). The DNB members and the correlations among these molecules were visualized in molecular networks where node color corresponds to the standard deviation and edge width corresponds to the correlation (Figure 4B). Remarkably, the DNB members were highly fluctuated and were also highly correlated among themselves. However, DNB and all non-DNB species sparsely connected at NAS 3, which is a typical feature of the critical stage characterized by DNB. Furthermore, a majority of these DNB members in TAGs in plasma (37/39, Table 1) overlapped with those in liver tissue (37/64, Table 1), a sharp contrast to a lack of consistent changes of these TAGs ions in plasma and liver based on abundance (Figure 2A). These results suggest that it is possible to use only the DNB members of TAG in plasma to define the pre-NASH or critical state from steatosis to NASH. We performed similar analysis on other lipid species including SM, PC, and PE in plasma, as well as Cer, PS, and PE in liver tissue. The maximal CI values in all these species were also obtained at NAS 3 (Supplementary Figure S4). Further analysis of the total lipid species in DNB also confirmed this finding (Supplementary Figure S5). Taken together, DNB theory further validated the existence of the critical stage from steatosis to NASH when NAS reaches around 3, and the collective fluctuations of DNB lipids in plasma are closely correlated with those in the liver tissue at the critical stage just before the transition from hepatosteatosis to NASH.
Pathways and networks associated with TAG metabolism at the critical pre-NASH state

After identification of the critical stage from steatosis to NASH using lipidomic data and DNB theory, we next set out to investigate pathways and networks of lipid metabolism that may be responsible for this transition. TAG is the primary form of intracellular lipid storage and one of the most abundant and diverse lipid classes in the liver and plasma. We examined the expression patterns of genes directly associated with TAG synthesis and degradation in liver. The net amount of hepatic TAG is regulated by the synthesis and degradation pathways

Figure 2 Lipidomic analyses of mouse plasma and liver tissue. (A) Heat maps of the relative amount of TAG in mouse liver tissue and plasma. Each row represents an individual TAG and each column represents a NAS group. Color-coded rectangular cells are relative amount transformed to z-scores by subtracting the average of the mean abundance by absolute concentration and being divided by the standard deviation of the mean abundance of the six NAS groups. (B) Correlation of TAG between mouse liver tissue and plasma. (C) Correlation of all detected lipids between mouse liver tissue and plasma. PCC was calculated and indicated in y-axis. PCCs were grouped by NAS score and box-plotted. Student’s t-test without assuming equal variances was used to compare correlation difference between NAS < 3 and NAS > 3. (D and E) Hierarchical clustering of lipid ions in mouse liver tissue (D) and plasma (E) by their absolute amounts (n = 4–6 each group). Clearly, all of the samples were well clustered to their corresponded groups except the samples at NAS 3 (red color), which were scattered into different groups. This fact implies that the samples of NAS 3 are at a special state.
Acylation of DAG (diacylglycerol) to TAG is the final step of TAG synthesis, which is regulated by DGAT. On the other hand, TAG is converted to DAG by lipases including LPL and ATGL. We first examined the expression of these enzymes during the progression of NAFLD. Two isoforms of DGAT enzymes exist: DGAT1 and DGAT2. DGAT1 is expressed mainly in skeletal muscle, intestine, and skin, with lower levels of expression in liver and adipose tissues. DGAT2 is the major isoform in hepatocytes and adipocytes (lipid droplets). Gene expression of DGAT2 increased significantly after NAS 3 whereas DGAT1 did not change significantly from NAS 0 to 5 (Figure 5B and Supplementary Figure S6A). Interestingly, however, LPL expression decreased significantly at NAS 4 while expression of ATGL increased dramatically after NAS 3. These data appear to suggest that TAG accumulation in the liver (NAS > 3) is the net outcome of DGAT2 overexpression and synergistic down-regulation of LPL. Similar to the DNB theory, we calculated the coefficient of variation (CV) of these genes in an attempt to catch the fluctuations of pathways that led to the critical transition from steatosis to NASH. As shown in Figure 5C, an increase in CV occurred at NAS 2 both for ATGL and LPL, preceding the transition that we identified at NAS 3. Thus, hydrolysis of TAG by LPL appears to be critically important in controlling the hepatic TAG accumulation and critical transition from steatosis to NASH (NAS 3). This is a reminiscent of a recent discovery of critical genes that control TAG degradation in NAFLD pathogenesis including PNPLA3 (vide infra). We also studied other genes associated with DAG synthesis and breakdown, fatty acid de novo

**Figure 3** A brief description of DNB theory. (A) Representative pictures of H&E staining for the three states of NAFLD. Clearly, there is no significant difference between steatosis and pre-NASH but significant difference between NASH and pre-NASH. (B) Three criteria for DNB molecules and the quantitative index $C_{I_d}$ are depicted. Based on average expressions of lipids, pre-NASH state is indistinguishable from steatosis but is different from NASH. Thus, the traditional biomarkers based on differential expressions failed to identify pre-NASH samples from steatosis. However, DNB method can identify pre-NASH samples based on the three criteria due to differential correlations and differential deviations rather than differential expressions. (C) Network view of the DNB molecules and alterations of their interactions at different disease stages. At the pre-NASH state, DNB is the network satisfying the three criteria. Edge width corresponds to the correlation, and node color corresponds to the standard deviation of the nodes.
A transition state from hepatosteatosis to NASH

In this study, we identified a critical stage (pre-NASH) just before the transition from hepatosteatosis to NASH in a mouse model of high-fat-induced NAFLD using lipidomics and DNB. For the first time, lipidomic data were combined with DNB theory to identify key lipid species especially for TAG and their network at the critical stage. We also observed a sudden loss of correlations among several lipid classes and fatty acids in circulation with hepatic tissue, highlighting the challenge of using lipid biomarkers in the blood to track and reflect the disease progression in the liver, especially during the transition from steatosis to NASH. Furthermore, our newly discovered DNB members of TAG in the plasma were well overlapped with those in liver tissue, implicating their potential to serve as novel biomarkers to detect the early-warning signals of NASH or differentiate pre-NASH from steatosis, especially when the clinically used NAS falls in the borderlines.

Since diagnosing or staging NAFLD remains a clinical challenge, numerous efforts have been made to discover convenient and accurate biomarkers of NAFLD in the blood. Individual or combination of routine markers such as ALT, AST, oxidative stress markers as well as inflammation factors have been tested for this purpose but only achieved limited success. The underlying mechanisms responsible for these challenges remain to be understood. For the past decade, high expectation has been given to various omic techniques including genomics, proteomics, and metabolomics that generate high-throughput data to investigate pathogenesis of NAFLD at a system level (Younossi et al., 2005). With the rapid development of modern MS technology, lipidomics emerges as a tool of choice to discover potential lipid biomarkers that can be used to diagnose NAFLD (Lamaziere et al., 2012). Recent technological advances in lipidomics enable researchers to measure hundreds of lipid species routinely (Han et al., 2012). Emerging lipidomic studies demonstrate that alterations of critical lipid classes, such as glycerolipid, phospholipid and sphingolipid, are strongly associated with NAFLD. A recent study by Gorden et al. (2015) proposed that a panel of 20 plasma metabolites including glycerophospholipids, sphingolipids, sterols, and various aqueous metabolites involved in cellular metabolic pathways could be used to differentiate NASH from steatosis. While discrimination between those two states requires broader panels of metabolites for efficient classification, this study highlights the challenges and complexity of lipid metabolism during progression of NAFLD (Gorden et al., 2015). Furthermore, it is still a long way to go before any lipid biomarkers can be used in clinical diagnosis and classification of NAFLD.

System-wide analysis based on omic data has been widely used to identify novel disease markers (Wu et al., 2014). It is of great importance to discover biomarkers in the blood that can reflect the existing or progression of NAFLD in the liver. However, only limited success has been reported in the field and the mechanism underlying this challenge remains elusive. Conventional disease biomarkers are based on the differential expressions of metabolites in the disease stage compared with normal state. It is conceivable that the underlying biological pathways reflecting a specific disease biomarker are dynamic and functionally changed during the disease development, progression or treatment. This point is evident in our current lipidomic study where the correlations among different lipid species and fatty acids suddenly dropped at the critical stage or pre-NASH stage (NAS 3) (Figure 2B and C). This is a significant observation that highlights the challenges in finding blood lipid markers to track the disease progression based on the conventional disease biomarkers. This loss of correlation may be due to differences in lipid species turnover rates between plasma and tissue compartments and due to the fact

Note: DNB members of TAG in both plasma and liver were indicated in bold. The TAGs shown are the major species based on the relative distribution of fatty acids.

<table>
<thead>
<tr>
<th>Table 1 The DNB of TAG in the plasma and liver.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plasma DNB of TAG</strong></td>
</tr>
<tr>
<td>TAG C5:0</td>
</tr>
<tr>
<td>TAG C5:3</td>
</tr>
<tr>
<td>TAG C5:2</td>
</tr>
<tr>
<td>TAG C5:1</td>
</tr>
<tr>
<td>TAG C5:0</td>
</tr>
<tr>
<td>TAG C5:1</td>
</tr>
<tr>
<td>TAG C5:0</td>
</tr>
<tr>
<td>TAG C5:3</td>
</tr>
<tr>
<td><strong>Liver DNB of TAG</strong></td>
</tr>
<tr>
<td>TAG C4:3</td>
</tr>
<tr>
<td>TAG C4:2</td>
</tr>
<tr>
<td>TAG C4:1</td>
</tr>
<tr>
<td>TAG C5:0</td>
</tr>
<tr>
<td>TAG C5:1</td>
</tr>
<tr>
<td>TAG C5:2</td>
</tr>
<tr>
<td>TAG C5:3</td>
</tr>
<tr>
<td>TAG C5:4</td>
</tr>
</tbody>
</table>

**Note:** DNB members of TAG in both plasma and liver were indicated in bold. The TAGs shown are the major species based on the relative distribution of fatty acids.
that the plasma lipid profile is a composition of lipid derived from different compartments. How a certain disease state exacerbates the correlations between circulation and certain tissue compartment has never been reported. There is only one closely related study that appeared in a recent literature, which found that plasma acylcarnitines inadequately reflected tissue acylcarnitine metabolism by examining the correlation of plasma acylcarnitine levels with selected tissue acylcarnitines in both fed and fasted BALB/cJ and C57BL/6 mice (Schooneman et al., 2014). However, this study was not carried out in the context of disease progression.

Due to the aforementioned limitations in characterizing disease stages simply based on absolute abundances of molecules using conventional biomarkers, we recently developed the DNB method to detect the critical stage just before the transition from normal to disease state during the progression of complex diseases by exploiting deviation and correlation information of metabolites based on nonlinear dynamical theory. The specific behaviors of the DNB members at a critical stage are the earliest signpost of deterioration. The abnormality of a group of molecules in this state may drive the disease progression but this process can be held back if a proper intervention is carried out. The DNB members differ from traditional molecular and static biomarkers and may provide further mechanistic insights into disease progression. Our work represents the first attempt to differentiate pre-NASH from steatosis by lipidomics. We identified DNB members of multiple lipid species during NAFLD progression when NAS approaches 3, coinciding with current challenges in differentiating steatosis from NASH (in particular pre-NASH) even based on liver biopsy (Figure 4). More interestingly, majority of the TAG DNB molecules in the blood overlapped with those in the liver tissue (Table 1), a stark contrast to a poor correlation of lipid species in the blood and liver tissue based on abundance (Figure 2D and E). Thus, our observation may represent a

Figure 4 Identification of the DNB of TAG lipids in mouse plasma and liver of HFD-induced NAFLD. (A) Quantitative index CI of the DNB of TAG in mouse plasma and liver. (B) Dynamic behavior of DNB members in TAG of mouse plasma and liver, respectively. Upper small circles represent identified DNB members and lower circles are non-DNB members. Node color reflects the standard deviation of the corresponding lipid ions. The strength of correlations is reflected by edge width, where a wider edge corresponds to a higher correlation. For clear views, connections between non-DNB members were not plotted and the nodes were filled with gray color.
A transition state from hepatosteatosis to NASH

Figure 5 Pathways and dynamical behavior of gene expression associated with TAG metabolism. (A) Pathways of TAG metabolism and associated enzymes. (B) Gene expression of DGAT2, ATGL, and LPL in mouse liver. The results are presented as mean ± SEM of gene expression fold change. 36B4 gene is used as internal standard. *P < 0.05 comparing each NAS group with NAS 0 (n = 4–6 for each group). (C) Coefficient of variation (CV) of liver DGAT2, LPL, and ATGL gene in the progression of NAFLD.

fundamental phenomenon that perturbation of metabolic pathways in a specific tissue can be readily reflected in the blood. If validated in human samples and other diseases, it would have a profound impact to challenge the current paradigm of discovery of disease biomarkers. We are currently exploring this concept to use DNB as potential disease biomarkers for staging disease progression, even for identification of a ‘pre-disease’ state.

According to the DNB theory, the molecules exhibiting DNB properties not only can be used as biomarkers but also are strongly related to ‘drivers’ of the disease progression. In current study, we attempted to examine the underlying metabolic pathways associated with the identified DNB molecules, especially TAGs. We primarily focused on the pathways of TAG synthesis and degradation (Figure 5A). Genes associated with these pathways have been well studied and implicated in NAFLD pathogenesis (Schweitzer and Finck, 2014). However, no study has ever examined variation and correlations of their expression in the context of NAFLD progression. Previous studies reported that reduction of DGAT2 by antisense oligonucleotide improved hepatic steatosis and hyperlipidemia in obese mice (Yu et al., 2005); lipoprotein lipase expression in livers of obese subjects is an important factor responsible for liver steatosis (Pardina and Peinado-Onsurbe, 2012). In our
In the present study, we confirmed the patterns of gene expression as previously reported and further identified that collective fluctuations of TAG hydrolysis enzyme LPL and ATGL preceded the critical transition from steatosis to NASH transition. Thus degradation of TAGs may represent an inherent mechanism that is highly critical during the NAFLD progression. Our study is consistent with human GWAS study that identified PNPLA3 as the most important gene modifier in all aspects of the NAFLD spectrum (Anstee and Day, 2013; Smagris et al., 2015).

In summary, we discovered a critical stage from nonalcoholic hepatosteatosis to NASH in a mouse model using lipidomic technology and the DNB theory. This transition occurs after NAS 3, a critical point where current diagnostic method often fails to detect even based on the gold standard of liver biopsy. This critical stage also features sudden loss of correlations of all lipid species including total fatty acids in the blood and liver tissue based on the abundances. This observation seriously challenges our current paradigm to discover lipid biomarkers in the blood to intimately track the progression of NAFLD. However, our DNB method has advantages in this regard because almost all of the DNB molecules of TAGs in the blood overlap with those in the liver tissue. In addition, in contrast to the traditional methods or biomarkers based on differential expressions of molecules, our DNB method exploits differential correlations and differential deviations of molecules, thereby diagnosing the disease state in an accurate manner. We further investigated the potential mechanisms attributed to this critical transition and pathways associated with TAG synthesis and degradation appeared to be critical, especially for the enzymes that are responsible for TAG hydrolysis. Since metabolomic and lipidomic data are generally more consistent between humans and animal models compared with genomics and proteomics, it is conceivable that the discovered DNB members at this critical stage in mouse model of NAFLD can be validated in human NAFLD. If validated in humans, it would have a profound impact on the current paradigm in NAFLD diagnosis and management.

Materials and methods

Chemicals

Solvents for sample preparation and mass spectrometry analysis, such as methanol (MeOH), chloroform (CHCl3), and water, were purchased from Burdick and Jackson. All other chemical reagents were from Sigma-Aldrich. A detailed list of all lipid standards spiked into each sample can be found in Supplementary Table S1.

Animals

All animal experiments were conducted in accordance with protocols approved by Institutional Animal Care and Use Committee of the Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. C57BL/6 male mice (purchased from SLRC, Shanghai, China) at 6 weeks of age were fed either a regular chow diet (70 mice) or HFD (60% fat, D12492, Research Diets, New Brunswick, NJ) (110 mice) ad libitum for 6–18 weeks and were sacrificed at the designated time points. The experiments were carried out in two batches and similar results were obtained for each batch. Body weight was measured weekly. After overnight fasting, mice were anesthetized by intraperitoneal injection of chloral hydrate. Tissue of the liver from the same position of the largest lobe was cut and fixed in 10% neutral buffered formalin for histological examination and the remaining part was snap-frozen in liquid nitrogen and stored at −80°C for measurement of hepatic lipids and mRNA extraction. Blood samples were obtained simultaneously in a K2-EDTA containing tube and plasma was separated after centrifugation with 8000 g at 4°C for 5 min. The plasma samples were stored at −80°C for lipidomic analysis. For liver histological examination, H&E staining was performed and histological NAS was assessed by an experienced pathologist in a double blinded fashion. The scoring system comprised four histological evaluated semi-quantitative features: steatosis (0–3), lobular inflammation (0–2), hepatocellular ballooning (0–2), and fibrosis (0–4).

Measurement of metabolic parameters in plasma and liver

Plasma TAG, TC, ALT and AST were measured using assay kits (BHKT) as previously described (Du et al., 2012). Plasma glucose levels were measured by D-Glucose Enzymatic Bioanalysis Kit (R-Bio-pharm). For liver TAG, liver tissue was homogenized in PBS and extracted with CHCl3:MeOH (2:1) reagent and detected as plasma TAG by kit.

Preparation of lipid extracts from liver and plasma

All lipid extracts from mouse plasma and tissue samples were prepared as described previously with slight modification (Han et al., 2004). For liver tissue, a liver sample (10–20 mg) was weighed and homogenized in 200 μl of ice cold and filtered 10 × PBS (1:10 dilution) buffer. After determination of protein concentration by BCA assay (Pierce), an internal standard mixture for lipid analysis was added to each homogenate according to protein concentration. For plasma lipiddomics, 200 μl of plasma was extracted after addition of internal standard mixture. To improve specificity and sensitivity of PE detection, a derivatization method was used as previously published (Han et al., 2005).

Electrospray ionization-mass spectrometry-based shotgun lipiddomics

The shotgun lipiddomics was performed on a TSQ Vantage triple quadrapole mass spectrometer (Thermo Fisher Scientific) according to published protocols. Tandem MS scan fragment and collision energy for each lipid class were optimized as previously reported (Yang et al., 2009) and were listed in Supplementary Table S2. For heatmap presentation, relative amount of each lipid ions was used in Figure 2A while absolute lipid amount of each lipid species in every mouse was used for Figure 2B–E.

DNB analysis

Detailed theoretical description of DNB was reported previously (Chen et al., 2012). In brief, DNB is a group of molecules satisfying the following three requirements only at the critical state:

1. The markers or molecules are highly fluctuated, i.e. their average standard deviation of the markers (DNB members) is large.
2. The markers are highly correlated, i.e., their average correlation among all pairs of the markers (DNB members) is high.

3. The markers are ‘isolated’ from other molecules, i.e., the average correlation between markers (DNB members) and all non-markers (all non-DNB members) is low.

Based on nonlinear dynamical theory, whenever DNB with these above three features appears, the system is at the critical stage, i.e., with the further progression of a disease, the system will undergo drastic deterioration from the normal state to the disease state (Figure 3B). The state at this critical stage is the pre-disease state, or pre-NASH state in this work. Clearly, the pre-NASH state is still one part of the steatosis state but is its critical state just before the transition from steatosis to NASH (Figure 3B). For a given stage or time point, we designed a quantitative measurement, CI, to quantitatively characterize the extent to which the three properties are satisfied for a group of specific molecules, i.e., a set $d$,

$$\text{CI}_d = \frac{\text{SD}_d \cdot \text{PCC}_d}{\text{PCC}_d},$$

(1)

where $\text{SD}_d$ and $\text{PCC}_d$ are average standard deviation and average PCC of all molecules in $d$, while $\text{PCC}_o$ are the average correlation between molecules in $d$ and others that are not in $d$. Suppose $X^k = (x^k_i)$ is the raw data matrix in stage $k$, where $i$ represents the lipid species (or feature) and $j$ the sample number. To identify the DNB, we need to find the molecular set $d$ that has the highest value of $\text{CI}_d$. To do this, we developed the following computational algorithm:

1. Data quality control
   (a) For a given feature or molecule, outliers are replaced by overall mean of that feature. Here the outlier is defined by $x^k_i$ where $x^k_i = q_1 - 3(q_3 - q_1)$ or $x^k_i \geq q_3 + 3(q_3 - q_1)$ where $q_1$ and $q_3$ are 25% and 75% percentiles of feature $i$, respectively.
   (b) Data are normalized along stage or time point, i.e. suppose $X^h$ is the normalized data matrix of $X^k$, then $Y^h = (X^h - \mu) / \sigma$, where $\mu$ and $\sigma$ are the overall mean and standard deviation of $X^h$.

2. Identifying DNB
   (a) Collect features, of which the standard deviations are above the 50% percentile or the features that are significantly different ($P$-value of $t$-test $<0.05$) between case and control groups. Denote this collected feature set as $C$.
   (b) Hierarchically cluster $C$ to get modules (i.e. sets of molecules that are clustered together) of size $\geq 5$. The distance is defined as $1 - |\text{PCC}|$ and the cutoff is set to 0.2. Denote the resulting modules as $D = \{d\}$.
   (c) For any module $d \in D$, calculate the $\text{CI}_d$.
   (d) Find the module $d^*$ with maximum index, i.e. $d^* = \arg \max_{d \in D} \text{CI}_d$. We call $d^*$ as a responsive module. Find the module with maximum CI among all responsive modules. This module is most likely to be DNB and the corresponding stage is the critical stage, at which the system is at the pre-disease state.

Furthermore, we performed the Fisher’s $z$-transformation to PPCs and obtained the similar results except that the CI scores are slightly different from those derived from the raw PCCs without the normalization.

Generally, although there are no significantly differential expressions between steatosis and pre-NASH states, there are significantly differential correlations and differential deviations between those two states. Thus, DNB method can detect the pre-NASH stage.

Gene expression by real-time quantitative polymerase chain reaction

Total tissue RNA was extracted with TRIzol® (Invitrogen) and the concentration was determined by spectrophotometry. RNA (1 μg) was subjected to reverse transcription PCR by TaKaRa PrimeScript® RT reagent Kit. Diluted cDNA was then subjected to SYBR Green-based real-time quantitative polymerase chain reaction analysis (Roche FastStart Universal SYBR Green Master (ROX)). Primers were listed in Supplementary Table S3.

Statistics

Data are presented as mean ± SEM. Differences between group means were analyzed by 2-tailed, unequal Student $t$ tests using GraphPad Prism 5.0 (GraphPad Software). $P$-values $<0.05$ were considered statistically significant.

Supplementary material

Supplementary material is available at Journal of Molecular Cell Biology online.

Acknowledgements

We would like to thank Drs Shengxian Li and Xiaoyu Chen at Renji Hospital affiliated with Shanghai Jiao Tong University for helping with histological assessment of liver samples. We also acknowledge the critical review of this manuscript by Dr Xiaoying Li at Ruijin Hospital affiliated with Shanghai Jiao Tong University.

Funding

This work was supported by grants from the National Key Basic Research Program of China (973 Program, 2012CB524905, 2014CB910504), the Strategic Priority Research Program of the Chinese Academy of Sciences (CAS) (XDB13040700), the National Natural Science Foundation of China (31170809, 91439103, 31470831, 61134013, 91529303, 81471047), the Ministry of Science and Technology of China (2012BAK01B00, 2013BAI04B03), and the Hundred Talents Program from CAS (2012OHTP07). H.Y. is an Associate Fellow at the Collaborative Innovation Center for Cardiovascular Disease Translational Medicine at Nanjing Medical University.

Conflict of interest: none declared.

References


