Detecting tissue-specific early warning signals for complex diseases based on dynamical network biomarkers: study of type 2 diabetes by cross-tissue analysis

Meiyi Li*, Tao Zeng*, Rui Liu and Luonan Chen

Submitted: 12th February 2013; Received (in revised form): 25th March 2013

Abstract

Identifying early warning signals of critical transitions during disease progression is a key to achieving early diagnosis of complex diseases. By exploiting rich information of high-throughput data, a novel model-free method has been developed to detect early warning signals of diseases. Its theoretical foundation is based on dynamical network biomarker (DNB), which is also called as the driver (or leading) network of the disease because components or molecules in DNB actually drive the whole system from one state (e.g. normal state) to another (e.g. disease state). In this article, we first reviewed the concept and main results of DNB theory, and then applied the new method to the analysis of type 2 diabetes mellitus (T2DM). Specifically, based on the temporal-spatial gene expression data of T2DM, we identified tissue-specific DNBs corresponding to the critical transitions occurring in liver, adipose and muscle during T2DM development and progression. Actually, we found that there are two different critical states during T2DM development characterized as responses to insulin resistance and serious inflammation, respectively. Interestingly, a new T2DM-associated function, i.e. steroid hormone biosynthesis, was discovered, and those related genes were significantly dysregulated in liver and adipose at the first critical transition during T2DM deterioration. Moreover, the dysfunction of genes related to responding hormone was also detected in muscle at the similar period. Based on the functional and network analysis on pathogenic molecular mechanism of T2DM, we showed that most of DNB genes, in particular the core ones, tended to be located at the upstream of biological pathways, which implied that DNB genes act as the causal factors rather than the consequence to drive the downstream molecules to change their transcriptional activities. This also validated our theoretical prediction of DNB as the driver network. As shown in this study, DNB can not only signal the emergence of the critical transitions for early diagnosis of diseases, but can also provide the causal network of the transitions for revealing molecular mechanisms of disease initiation and progression at a network level.

Keywords: Type-2 diabetes; disease progression; critical transition; dynamical network biomarker (DNB); leading network; multi-tissues

*These authors contributed equally to this work.

Meiyi Li is a PhD student at Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. Her main interest is in developing computational methods to understand disease progression at a molecular level.

Tao Zeng is an Assistant Professor at Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. His research interest includes network biology and computational biology.

Rui Liu is an Assistant Professor at Department of Mathematics, South China University of Technology, China. His main interest includes computational systems biology and non-linear analysis on biological systems.

Luonan Chen is Professor and Executive Director of Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. His research interest is computational systems biology and bioinformatics.
INTRODUCTION

Recent rapid advance in high-throughput technologies has greatly expanded the availability of information at all levels of biological systems [1, 2], which has significantly enhanced the research of translational medicine [3, 4], in particular on discovering new biomarkers of complex diseases, e.g. cancers and diabetes. Considerable work has been done on the biomarkers for disease diagnosis or prognosis [5–12]. These biomarkers or disease models mainly measure disease status of patients or provide the necessary information for clinical judgments of disease states [13, 14]. But, they usually cannot be directly used to evaluate or predict a person in a pre-disease state (e.g. the state before the appearance of disease symptoms), which is the key for achieving early diagnosis of diseases.

Generally, a disease progression can be divided into three stages (Figure 1A), i.e. normal state (a stable state where the system undergoes gradual or slow change), pre-disease state (a limit of the normal state just before the drastic transition to the disease) and disease state (another stable state after the critical transition, where the system is considered to be seriously damaged) [15, 16]. Theoretically, the pre-disease state corresponds to a critical point before the system transits to an irreversible disease state accompanying a drastic change of system dynamics [17–19]. Thus, it is crucial to detect the pre-disease state so as to achieve the early diagnosis of complex diseases. Traditional biomarkers are able to distinguish disease state from normal state mainly based on the differential expressions on individual molecules or a group of molecules, but are unable to detect the pre-disease state owing to their static nature. Completely different from the traditional methods, recently, a novel method based on dynamical network biomarker (DNB) has been developed to distinguish the pre-

![Figure 1](https://academic.oup.com/bib/article-abstract/15/2/229/211863)
disease state from the normal state. On the basis of non-linear dynamical theory, DNB is able to detect the early warning signals of the critical transitions during the disease progression, by exploiting rich information from the observable high-throughput data [20, 21]. As shown in Figure 1B, in contrast to the static information of molecular expressions mainly adopted in traditional methods or biomarkers, DNB explores the dynamical information of molecular fluctuations as well as the correlations between molecules, which is a new concept on biomarkers from both theoretical and biomedical viewpoints. This method has been validated by theoretical analysis and experimental data on various diseases [17, 18]. In addition to the early diagnosis of diseases, DNB can be also applied to detect key state changes and their driving networks during many biological processes, e.g. cell differentiation processes and proliferation processes [17, 19].

On the other hand, the globe figure of people with diabetes is increasing rapidly [22], especially for type 2 diabetes mellitus (T2DM), which is a chronic disease with nature history lasting for >20 years. Despite intensive studies on T2DM, its molecular mechanism and its early molecular markers still remain unclear. Generally, T2DM can be divided into five stages: latent stage, transition stage, impaired glucose tolerance (IGT) stage, impaired fasting glucose stage and overt stage [23]. During the first four stages, the sub-health status of patient is still able to return to normal, and IGT and impaired fasting glucose stages are called pre-diabetes. Whereas, T2DM is always diagnosed when reaching the overt stage. Therefore, it is a key problem to identify when and how the transition between different stages happens at a molecular level, in particular, ahead of the diagnosis of T2DM [24–29]. Clearly, the early warning signals of critical transitions during T2DM progression can be used for not only achieving prevention and personalized medicine, but also revealing molecular mechanisms of the disease development.

T2DM is a well-known metabolic disease characterized by insulin resistance and beta cells failing to compensate, so that insulin-responsive tissues are important organ targets of T2DM development and progression [24–29]. Those tissues include muscle, liver and adipose, each of which actually has its own characteristics in disease progression. As T2DM (also known as non-insulin-dependent diabetes) attributes to the metabolic disorders mainly induced by the complicated interplay of both genetic factors and environmental factors [30], there are intensive system-wide studies on T2DM [26, 28, 29, 31–33], e.g. GWAS. And as a result, many genetic and epigenetic factors involved in the disease progression at different tissues have been identified [22, 34]. Despite these progresses, however, many questions remain unanswered, i.e. (i) the critical transition points on each insulin-responsive tissue are still unclear, (ii) the driver molecules of those transitions that result in T2DM have not yet been fully revealed and (iii) the dynamical roles of these insulin-responsive tissues in T2DM development and progression have not yet been clearly understood, especially in the presence of pre-disease state. To answer these questions, in this study, based on multi-tissue gene expression data of a T2DM animal model, we identified the critical transitions and their driver (or leading) networks on muscle, adipose and liver by tissue-specific DNBs and further analyzed the dynamical regulations during T2DM development and progression by both intra-tissue and inter-tissue studies. Specifically, in this article, we first gave a brief summary of the theoretical results of DNB and its computational algorithm [15, 16] for detecting early warning signals of T2DM; then analyzed the dynamical behaviors of early warning signals in each insulin-responsive tissue; next, discussed the biological significance (T2DM relevance) of tissuespecific DNBs from various aspects, such as single disease genes, functional enriched gene sets and conditional rewired gene sub-networks; furthermore, investigated the spatial-temporal characteristics of DNBs and their relations with T2DM development and progression and finally, concluded this study by providing several general remarks.

**METHODOLOGY OVERVIEW**

**Theoretical foundation on DNB**

As shown in Figure 1A, we consider that the biological process of a transition from normal to disease (or from early disease to advanced disease) can be divided into three stages [15], i.e. ‘before-transition state’ (or normal state), which is a stable state and where the system undergoes gradual or slow changes; ‘pre-transition state’ (or pre-disease state), which is a limit of the ‘before-transition state’ just before the critical transition to another state and ‘after-transition state’ (or disease state), which is another stable state. When a biological system transfers from one state to another one during disease development and
progression, some relevant molecules integrating into a network of this system will experience a drastic transition, which drives the corresponding system to have a qualitative change.

Then following the derivation of [15], we consider that the state dynamics of a living organism can be described by a non-linear dynamical system Equation (1),

$$Z(k+1) = f(Z(k); P)$$

where the vector $Z(k) = (z_1(k), \ldots, z_d(k))$ represents observed data with $n$ observed variables, i.e. molecule concentrations (e.g. gene expressions or protein expressions) at time point $k$ ($k = 0, 1, \ldots$), e.g. minutes or days, and it is the variables indicating the dynamical state of a biological system. Parameter $P$ represents slowly varying factors such as genetic factors (e.g. SNP and CNV) and epigenetic factors (e.g. methylation and acetylation), which actually represents slowly varying factors such as genetic conditions.

Parameter $P$ is the dynamical state of a biological system. Parameter $P$ is unknown parameters, which are not required to be measured in this study, and the dynamics of $P$ are considered much slower than $Z$.

As the system is stable at the normal state (a stable fixed point $\mathcal{Z}$), we consider the linearized equations of Equation (1) at $\mathcal{Z}$. For simplicity, we assume that all eigenvalues of the Jacobian matrix of $f$ at $\mathcal{Z}$ are real values (for other cases, see [6, 15]). Then, there exists a transformation matrix $S$ so that the linearized Equation (1) can be transformed to the following Equation (2):

$$Y(k+1) = \Lambda(P)Y(k) + \xi(k)$$

where new variable $Y(k) = (y_1(k), \ldots, y_m(k))$ corresponds to the original variable $Z(k)$, i.e. $Y(k) = S^{-1}(Z(k) - \mathcal{Z})$, $\Lambda(P)$ is the diagonalized matrix of $\frac{\partial f(Z(k); P)}{\partial Z}|_{Z=\mathcal{Z}}$ and noise vector $\xi(k) = (\xi_1(k), \ldots, \xi_m(k))$ represents Gaussian noises with zero means and covariances $\kappa_{ij} = \text{Cov}(\xi_i, \xi_j)$.

Let $\lambda_1$ be the largest eigenvalue, and then at the normal state, $|\lambda_1| < 1$. When the system approaches the critical point, $|\lambda_1| \to 1$. Then, those molecules, $z_i$ (e.g. genes, proteins or small compounds), affected by $\lambda_1$ form a dominant group of molecules, which has been proven to satisfy the following three conditions:

1. Concentrations of molecules ($z_i$) in the dominant group for pre-disease samples strongly fluctuate comparing with those for normal samples. In other words, the average SD of molecule concentrations (denoted by $SD_d$) is significantly higher in the group.

2. Molecules in this group become closely related to each other, which means that their average Pearson correlation coefficient (PCC) in absolute value between all $z_i$ and $z_j$ (denoted by PCC$_{ij}$) in the group drastically increases, compared with the value for normal samples.

3. Oppositely, the average PCC between molecules in this group and others decreases drastically, i.e. the corresponding average PCC of molecules in absolute value (denoted by PCC$_{ij}$) becomes smaller, contrary to the value for normal samples.

If there is one group of molecules satisfying all three conditions simultaneously, it is the dominant group, or so-called DNB, whose existence implies the imminent transition from one stable state to another. As shown in Figure 1B, in contrast to the conventional molecular biomarkers distinguishing disease samples from normal samples by mainly using molecular expression information (i.e. static information), the DNB distinguishes not disease samples but pre-disease samples from normal samples by using molecular fluctuation information (i.e. dynamical information) and also network information (i.e. correlation information among molecules). In addition, DNB is a model-free approach, which requires no settings on parameters or even models. In other words, patients may have different DNBs, even suffering from the same disease due to the personal variations, and thus this method is a general measurement on diseases. Therefore, one advantage of DNB can be expected to directly apply to the personalized medicine. Note that theoretically there exist at most two DNBs for one critical transition [15, 16], either of which can be used to signal the imminent transition. For a disease process, there may be multiple transitions resulting in different deterioration stages.

**Algorithm for detecting early warning signal of a critical transition**

The dominant group characterizes dynamical features of the underlying biological system and the molecules in the group are also strongly and dynamically correlated in the critical transition period (e.g. pre-disease state); the molecules in the dominant group are expected to form a sub-network of the biological system, which is a dynamical network of biomarkers,
or a DNB. DNB is also called as the driver (or leading) network because the members of the DNB make the first move from the one state to another. As stated in the three conditions, the presence of DNB can be observed by the measurement of dynamics or multiple samples. DNB appears in the critical transition period but disappears in other periods, thereby signaling the upcoming critical transition.

To obtain a strong signal at the critical transition, the three criteria were further combined together to construct a composite index (CI) as follows:

\[ CI = \frac{SD_d \times PCC_d}{PCC_o} \]  

where \( SD_d \) is the average SD of the molecules of the dominant group or DNB, \( PCC_d \) is the average PCC of the DNB in absolute value and \( PCC_o \) is the average PCC of molecules between those in DNB and other molecules in absolute value. The CI is expected to increase sharply (i.e. reaches maximum) when the biological system approaches a critical transition or pre-disease state. Therefore, CI can indeed serve as an early warning signal to identify the pre-disease state effectively, although other index can be also constructed, provided that the three conditions are satisfied.

The detailed flowchart of computational algorithm for detecting DNB according to CI is showed in Figure 2 and can be also stated as follows:

1. Deviation test: select the genes with significantly high deviations at each period or state based

![Figure 2: Algorithm for detecting early warning signal of a critical transition.](https://academic.oup.com/bib/article-abstract/15/2/229/211863)
on expression profiles (e.g. gene expressions or protein expressions for a number of samples). Then, we have a group of molecules with high SDs at each period $t$, i.e. a set of molecules $N_{tD}$;

(2) Intra-correlation test: cluster the previously chosen genes from $N_{tD}$ at each period or state based on correlations. Then, we have a smaller group whose molecules have high correlations at each period $t$, i.e. $N_{tI}$ which is a subset of $N_{tD}$;

(3) Inter-correlation test: choose the genes from $N_{tI}$ whose correlations with other molecules outside of $N_{tI}$ are significantly low, at each period. Then, we have a smaller group $N_{tC}$ at each period $t$, which is a subset of $N_{tI}$;

(4) DNB test: determine the critical transition point and its corresponding dominant group (or module) or DNB from $N_{tC}$ ($t = 1, \ldots, T$) by CI significance analysis (e.g. CI score, or P-value), i.e. let $N_{tC}$ be DNB if its corresponding CI is the largest among all periods $t = 1, \ldots, T$;

(5) Functional analysis on DNB: analyze the enriched biological functions of the identified DNB, e.g. by GO or network ontology analysis [35] or pathway analysis [6, 36].

Based on the theoretical derivation [15, 16], both $SD_d$ and $PCC_d/PCC_s$ should be maximal at the transition stage or point. However, for many cases, we usually have observed data not at exact critical transition point but near the transition. Therefore, near the transition point, numerically, $SD_t$ and $PCC_d/PCC_s$ may not simultaneously reach maximum. In addition, to improve the computational efficiency or robustness, entropy or mutual information can be also adopted to measure associations between two molecules [16] instead of PCCs in the aforementioned computational algorithm.

RESULTS
Early warning signals of T2DM in multi-tissues
Here, we applied the method of DNB to detect the critical transitions during T2DM development and progression accompanying insulin resistance in adipose, gastrocnemius muscle and liver for rats [diabetes rats: Goto–Kakizaki (GK) rats; control rats: Wistar–Kyoto (WKY) rats]. The high-throughput experimental data sets were downloaded from NCBI GEO database (access ID: GSE13268, GSE13269 and GSE13270) (www.ncbi.nlm.nih.gov/geo), which were time-course gene expression data obtained from age-specific rats corresponding to well-designed time series (five samples in 4, 8, 12, 16 and 20 weeks, individually).

After performing the aforementioned algorithm for detecting early warning signal, we found two critical transitions for these tissues in the disease evolution of T2DM (Figure 3). As one of the most important energy metabolism systems, strong signal of CI for liver first appeared in rats at the age of 4 weeks (module L1) and then again at 16 weeks (module L4). For the muscle, it likely responded the signals from liver with the same critical transition periods at 4 and 16 weeks, respectively (i.e. module M1 and module M4). However, for adipose, there was only one critical transition period corresponding to the disease deterioration at the age of 8 weeks (module A2). Actually, these critical transition periods were also consistent with the development of T2DM in GK rats, according to the phenotype-changing trends of the plasma glucose and the plasma insulin of concentrations from GK and WKY rats at different ages (see Supplementary Figure S1). This phenotype information was provided by the authors [37] with the experimental data. Their studies described that the concentrations of plasma insulin in GK were sharply elevated between 4 and 8 weeks, but declined after 12 weeks, and finally were lower than WKY at 20 weeks. Meanwhile, the plasma glucose concentrations were presented at a higher level since 4 weeks of age and kept the increasing trend to 12 weeks [37]. These observations definitely revealed two different and important critical transition periods, one for insulin resistance and the other for beta-cell failure. This confirms our following results on dysfunction of insulin sensitivity and serious inflammatory response correspondingly. Especially, either in tissue-specific analysis or in cross-tissue analysis, a common functional cascade associated with T2DM has been found: (i) dysfunctions of insulin biosynthesis and response, (ii) resistance of plasma glucose uptake, (iii) energy supply, (iv) hormone response and (v) lipid metabolism that appeared to be abnormal or inflammatory.

Tissue-specific analysis
To confirm the significant relation between tissue-specific DNBs and T2DM progression, and to identify what signal could be provided to pre-warn the disease transitions or deteriorations, the GO analysis by network ontology analysis [35] and KEGG [38]
enrichments was conducted for each DNB to categorize the genes participating in different biological functions or pathways (Tables 1 and 2 and Supplementary Table S1).

Here, we first analyzed results from liver tissue, although both liver and muscle owned two common critical points. For GK rats at the age of 4 weeks, when the first critical point occurred, the genes of the identified DNB were mainly enriched in the basic metabolic processes and pathways, especially in the parts of lipid metabolic processes and some hormone and ion regulatory processes. However, at the second critical point (i.e. 16 weeks), the functions of this DNB mainly were enriched in the metabolism and transport of lipid and the inflammatory response. Specifically, in the first critical period, we found that the metabolism of xenobiotics by cytochrome P450 pathway played an important role in oxidation of organic substances, whose involved enzymes in liver could generally act as metabolic intermediates, e.g. lipids and steroid hormone [39, 40], whereas in the second critical period, the relevant biological processes and pathways for T2DM were enriched, including the regulation of interleukin, the regulation of PI3-kinase and calcium ion transport function [41]. In details, we further listed some interesting genes to explain how aforementioned functions are related to T2DM pathogenesis. The genes relevant to the first critical point, such as CYP11A1, CYP11B1, CYP11B2,

**Figure 3:** The composition index (A) and the DNB (B) for each tissue of GK rats. L1 and L4 are DNBs of liver at the first period (4 weeks) and the fourth period (16 weeks), respectively. M1 and M4 are DNBs of muscle at the first period (4 weeks) and the fourth period (16 weeks), respectively. A2 is the DNB of adipose at the second period (8 weeks). (C) visually illustrates dynamics of the identified DNBs compared with the whole molecular network.
Table 1: Functional enrichment of GO biological processes for each identified DNB

<table>
<thead>
<tr>
<th>Tissue</th>
<th>DNB</th>
<th>Genes in DNB</th>
<th>GO term</th>
<th>P-value</th>
<th>Term name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver</td>
<td>L1</td>
<td>ARHGEF9; SPTBN2; APC; CLASP2; CHGB; CYP1A1; CYP2T1; JUN; SHC3; AP; CHGB; CYP11A1; CYP2T1; JUN; SHC3; AP; CTNNB1; NFKB1; MARK3; DVL1; HSD3B; IL24; CTGF; SLC8A3; ...</td>
<td>0006694</td>
<td>3.60E-07</td>
<td>Steroid biosynthetic process</td>
</tr>
<tr>
<td>Liver</td>
<td>LI</td>
<td>PPP2R1B; LRPI; CAT; SPTBN2; APOA2; FABP4; RBP4; ITGA3; LRCIP1; CEBPA; CLU; ARHGEF9; EGR4; PHYH; HES3; PTBP3; CDC43; ACP6; ...</td>
<td>0022600</td>
<td>9.20E-04</td>
<td>Digestive system process</td>
</tr>
<tr>
<td>Muscle</td>
<td>MI</td>
<td>PPP2R2B; RXRA; PPARG; GCGR; NR3C1; MAPK13; APO2; FABP4; RBP4; ITGA3; LRCIP1; CEBPA; CLU; ARHGEF9; EGR4; PHYH; HES3; PTBP3; CDC43; ACP6; ...</td>
<td>0048511</td>
<td>1.70E-06</td>
<td>Rhythmic process</td>
</tr>
<tr>
<td>Muscle</td>
<td>M4</td>
<td>GNG13; FLT3; SLT2; CEBPA; PLDI; STAT5A; STAT5B; NFKB1; SCAFI; NRS1L2; STAT3; FLT3; CISH; SYNPR; CCR10; GABRG; TRAF3; SH2DA2; DMRT1; ...</td>
<td>0007259</td>
<td>7.30E-06</td>
<td>JAK-STAT cascade</td>
</tr>
<tr>
<td>Adipose</td>
<td>A2</td>
<td>PIP5K1B; CYPIIB2; GSTT2; AKTI1; AKTI2; AGPAT4; HSD11B2; AANAT; EGR1; GSTP1; PRDX3; PPAR2; CNR1; SMARCA1; CD9; OCLR; SGMS1; GMT7; ...</td>
<td>006644</td>
<td>9.00E-08</td>
<td>Phospholipid metabolic process</td>
</tr>
</tbody>
</table>

Li et al.
CYP21A1 and CYP2T1 as enzymes of the cytochrome P450 superfamily participate in many metabolic reactions of steroids, cortisol and other lipids, which is consistent with the fact that liver synthesizes and releases some very-low-density lipoprotein for T2DM [42]; CTNNB1 was also found to participate in the pathway of insulin internalization and SHC3 to participate in the insulin signal transduction, from the GeneCards (http://www.genecards.org/).

Moreover, owing to lacking adequate glucose uptake induced by dysfunction of insulin response, most pathways related to the cellular regulatory signal transduction of basic energy metabolism became abnormal, such as Wnt, MAPK and ErbB signaling pathways and so on. For genes involved in the second critical point, the kinase cascade reactions involved in PI3 played an important role in the metabolism of lipid, and calcium ion as an activator also participated in these reactions; PI3K–Akt signaling pathway was located at the upstream to regulate the inflammatory response, and this pathway was also identified by our method; FABP1 in liver could activate PPAR signaling pathway regulating gluconeogenesis and fatty acid oxidation [43] and also promote lipid transport through the function of PPAR–alpha; APOA2 associated with T2DM [44] played a key role in fatty acid transport as well as ITGAD in clearing lipoproteins. Interestingly, the retinol metabolisms have been found to play an important role in obesity and T2DM [45], in which some key genes (e.g. RBP4) were also detected in tissue-specific DNB.

For muscle at the onset of T2DM (or the pre-disease critical point), compared with liver at its first critical point, regulatory signaling pathways were enriched, and these pathways mainly regulate the system response to stimulus and inflammatory response. Some detected genes are related to the processes of regulating cAMP, Ca2+ and glucocorticoid, which are well-known regulators relevant to glycometabolism in basal metabolism. For muscle at the second critical point, responses to hormone stimulus and leukocyte activation were found in dysfunction, which were also observed in the first pre-disease critical point. At this latter critical point, lipid metabolisms were significantly abnormal too. However, the similar functions enriched in two critical periods of T2DM development and progression

<table>
<thead>
<tr>
<th>Tissue</th>
<th>DNB</th>
<th>KEGG term</th>
<th>Term name</th>
<th>Overlap</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 (liver 4 weeks)</td>
<td>00140</td>
<td>Steroid hormone biosynthesis</td>
<td>5</td>
<td>1.95E-08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04510</td>
<td>Focal adhesion</td>
<td>4</td>
<td>7.42E-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04300</td>
<td>Wnt signaling pathway</td>
<td>4</td>
<td>7.42E-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04010</td>
<td>MAPK signaling pathway</td>
<td>3</td>
<td>8.16E-03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04012</td>
<td>ErbB signaling pathway</td>
<td>3</td>
<td>1.12E-04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04146</td>
<td>Peroxisome</td>
<td>2</td>
<td>2.98E-05</td>
<td></td>
</tr>
<tr>
<td>Liver</td>
<td>00790</td>
<td>Folate biosynthesis</td>
<td>2</td>
<td>2.35E-06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>03230</td>
<td>PPAR signaling pathway</td>
<td>2</td>
<td>1.60E-03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04350</td>
<td>TGF-beta signaling pathway</td>
<td>1</td>
<td>1.63E-02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>00990</td>
<td>Oxidative phosphorylation</td>
<td>1</td>
<td>1.62E-02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04151</td>
<td>PI3K-Akt signaling pathway</td>
<td>8</td>
<td>3.51E-05</td>
<td></td>
</tr>
<tr>
<td>L4 (liver 16 weeks)</td>
<td>04010</td>
<td>MAPK signaling pathway</td>
<td>7</td>
<td>3.88E-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>00230</td>
<td>Purine metabolism</td>
<td>6</td>
<td>1.23E-05</td>
<td></td>
</tr>
<tr>
<td>Muscle</td>
<td>00240</td>
<td>Pyrimidine metabolism</td>
<td>5</td>
<td>3.93E-08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>03015</td>
<td>mRNA surveillance pathway</td>
<td>5</td>
<td>3.43E-06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>00564</td>
<td>Glycerophospholipid metabolism</td>
<td>5</td>
<td>2.16E-07</td>
<td></td>
</tr>
<tr>
<td>M1 (muscle 4 weeks)</td>
<td>04062</td>
<td>Chemokine signaling pathway</td>
<td>5</td>
<td>7.07E-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04630</td>
<td>Jak-STAT signaling pathway</td>
<td>4</td>
<td>2.18E-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04666</td>
<td>Fc gamma R-mediated phagocytosis</td>
<td>3</td>
<td>3.44E-04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04144</td>
<td>Endocytosis</td>
<td>3</td>
<td>7.62E-04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>00480</td>
<td>Glutathione metabolism</td>
<td>15</td>
<td>6.57E-23</td>
<td></td>
</tr>
<tr>
<td>M4 (muscle 16 weeks)</td>
<td>04666</td>
<td>Fc gamma R-mediated phagocytosis</td>
<td>10</td>
<td>5.78E-10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04062</td>
<td>Chemokine signaling pathway</td>
<td>10</td>
<td>1.57E-06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04151</td>
<td>PI3K-Akt signaling pathway</td>
<td>9</td>
<td>1.86E-03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>04144</td>
<td>Endocytosis</td>
<td>7</td>
<td>8.87E-06</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Functional enrichment of KEGG pathways for each identified DNB

Study of type 2 diabetes by cross-tissue analysis

237

Downloaded from https://academic.oup.com/bib/article-abstract/15/2/229/211863 by guest on 18 March 2019
actually resulted from different DNB genes (Supplementary Table S2). As a further example, the immune response mediated by those interleukins [regulated by several genes with sharply fluctuated expressions (e.g., NFKB1, STAT3, STAT5A, STAT5B) participating in the regulation of interleukin biosynthesis] played an important role in the second critical period, but the B cell receptor signaling pathway associating with PI3K–Akt regulation decided the immune response in the first critical period. This function conservation rather than gene conservation might be induced by decrease of insulin secretion and more severer resistance of glucose uptake (these clinical phenotypes will be discussed in details in next sub-section). In addition, there are also many such phenomena in muscle consistent with liver and adipose, for instance, a few genes (e.g. PGR, ESR1 and ESR2) related to responding hormone were also detected at 4 weeks, although steroid hormone biosynthesis does not belong to the muscle-specific functions. This fact might suggest that there are certain coordinated interactions or relations among those insulin-responsive tissues and the tissue-specific DNBs.

For adipose at its critical period, the enriched functions of its DNB mainly included lipid metabolisms and the response to stimulus. Similar to the dysfunctions of liver, the DNB genes of adipose had significant enrichments on the relevant biological processes and pathways of T2DM [41], such as the regulation of interleukin biosynthesis, dephosphorylation, phosphoinositide metabolism, the regulation of PI3-kinase, the activation of PKCs and calcium ion transport function. As a well-known T2DM-associated mechanism, PKCs activated by diacylglycerol made a contribution to the lipid-induced insulin resistance. In addition, steroid hormone biosynthesis as a non-specific function of adipose also became abnormal as observed in a liver-specific DNB.

Furthermore, to check the significance of our results by comparing the genes of aforementioned five DNBs with the genes (only 69 genes are included in this whole genomic expression profilers) reported in previous studies [46–50] (Supplementary Table S3), we found: RBP4 in DNB of liver at 16 weeks was identified to promote IGT (hypergeometric test, \( P < 2.56 \times 10^{-2} \)); 2 (GCGR and PPARG) of 58 DNB genes of muscle at the age of 4 weeks (hypergeometric test, \( P < 3.31 \times 10^{-3} \)) have been validated with significant relevance to T2DM; for adipose, there are also AKT2 of 146 genes (hypergeometric test, \( P < 0.13 \)) being reported as the disease gene. This supports again that DNB genes are key genes contributed to drive the critical transition during disease development and progression, which cannot be detected by conventional biomarkers and methods.

Cross-tissue analysis

From the aforementioned functional enrichment analysis, we have observed the comparative consistence among tissue-specific critical points, DNBs and common T2DM-relevant pathways. This means that there are cross-tissue functional relations between phenotype changes (e.g. blood glucose and insulin secretion) during the T2DM development and progression in GK rats, although these tissue-specific DNBs have few genes overlapped (Supplementary Table S2).

In the pre-disease period, the high blood glucose appeared in GK fed with the same food to two strains, although the insulin secretion level in GK rats was similar to in WKY rats (Supplementary Figure S1). Hence, it can be inferred that the insulin receptors in GK rat could not correctly respond to the insulin regulation, such that the glucose could not be transported successfully in cells. However, we have seen an interesting regulatory coordination on the biosynthesis and response to steroid hormone in aforementioned analysis. For liver and adipose, synthesizing and secreting steroid hormone had no tissue-specific functions, but they became abnormally active under this condition compared with WKY rats. Meanwhile, several genes participating in response to steroid hormone stimulus were also detected in muscle. In fact, the relation between steroids and insulin sensitivity has attracted increasing attention recently [51–53]. More importantly, for all tissues, the relevant function of fatty-acid transport was regulated through the PPAR signaling pathway activated by the upstream FABP family proteins, which might explain the significantly abnormal lipid metabolisms under the lack of glucose uptake.

In the pre-transition period, hyperglycemia still kept comparatively stable, but the concentration of plasma insulin was decreasing. According to the analysis of [24], the appearance of beta cells failing at 20 weeks might result from the increasingly serious lack of glucose uptake. Consequently, many biological processes involved in lipid metabolisms were activated in liver and muscle, such as lipid oxidation. And, many genes related to lipid transport became dysfunctional in liver. The previous studies showed
that, consistent to the T2DM sufferance, abnormal regulation of transporting with very-low-density lipoproteins was enriched in liver [42], and inflammatory response took place during the whole development of T2DM [54, 55]. However, the inflammatory response in liver and muscle tissues actually resulted from different regulations, i.e. B/T cell actively participated in the first critical period, and interleukins in the second critical period.

**DNB features**

To further focus on the specific characteristics of DNB as dynamical early warning signals, we compared DNBs with conventional biomarkers by their expression levels and roles in regulation. First, the DNB genes rarely belonged to differentially expressed genes (DEG) detected by T-test statistics with $P < 0.05$ at the same (critical) time point (Supplementary Table S4), and actually their overlapping significance measured by hypergeometric test was larger than 0.5. However, when checking DNB genes' expressions for all time points, it could be found that these genes had drastic over-expression or under-expression only at their corresponding critical points (fold-change $\geq 1.3$) (Supplementary Table S5). Second, we also found that many pathways related to T2DM contain genes from both DNB and DEG (Supplementary Table S4). This motivated us to assume that there may exist strong relationships between DNB and DEG as regulators and passengers. Actually, many DNB genes are really located at the upstream of the network to regulate DEG in a number of key biological pathways related to T2DM (Figure 4). Although DNB genes are not all in the upstream of DEGs, they stay at important places to regulate some branch metabolism paths. For instance, several DNB genes are located at the upstream of the pathway of steroid hormone biosynthesis, and many genes at the pathway downstream really present significant over-expression or down-expression. Some other DNB genes as messengers or receptors can help signals transduce into membrane, e.g. DNB genes participate in the upstream regulation of PI3K-AKT signaling transduction through GF–RTK interaction in muscle, whereas ECM affects ITG and FAK in adipose. Moreover, some genes in DNB were also found to act as a part of the combined complex to regulate the downstream dynamics, e.g. RXR with PPAR as a complex to regulate specific gene expressions in muscle at onset of T2D and G protein-gamma combine GPCR to regulate PI3K activation.

**DISCUSSION AND CONCLUSION**

In this article, we first briefly summarized the theoretical foundation of a novel network-based dynamical biomarker and its computational algorithm [15, 16] for detecting early warning signals of complex diseases, and then based on multi-tissue and multi-stage gene expression data during T2DM development and progression, we identified tissue-specific DNBs and further elucidated the role of DNBs as well as the molecular mechanism of T2DM at a network level.

The experiment data sets used in this work include the multi-period gene expression profiles of three different tissues from the typical diabetes rat model of the GK rats and the control model of the WKY rats. The three tissues are adipose, liver and gastrocnemius muscle, which are all major metabolic tissues with more risks to be affected by insulin resistance [30]. Based on this carefully designed temporal-spatial expression data, we showed that DNB method not only detected the early warning signals of stage transitions of each tissue in the development of T2DM but also identified the common biological functions or coordinated functions among multiple tissues. GK rat model for T2DM study has many advantages owing to its natural characteristics, such as high blood glucose, similar insulin resistance by the impaired insulin receptor and the multi-gene dysfunctions. Thus, due to the conserved genetic background and under similar condition, GK rats (relatively) synchronously suffer first from pre-disease symptom, then are in a disease state and finally reach the serious disease stage. GK rat model is also similar to the subset of type 2 diabetic human without obesity [37, 56]. Hence, this model is useful for understanding the common mechanisms of T2DM development for multiple tissues between human and rat.

In this work, we focused on liver, muscle and adipose (as three main target tissues of insulin) to find their tissue-specific DNBs and further analyze the molecular mechanism of T2DM at respective tissue during the disease progression. As one main result, the tissue-specific DNBs as early warning signals for T2DM development and progression were successfully detected. The results showed that DNB for each tissue and each transition point is actually different, thus justifying the necessity to conduct
tissue-specific analysis for complex diseases. Specifically, there are two critical states for both liver and muscle during T2DM development, the first (4 weeks) is involved in response to insulin resistance and the second (16 weeks) is related to the immune system. The identified DNB genes are significantly associated with T2DM, some of which have been reported to be the disease genes or

Figure 4: Key biological pathways with DNB genes and differentially expressed genes (DEG). Clearly, DNB genes are placed at the important positions in the pathways, i.e. core genes of the DNB are mainly located at the upstream to regulate DEG.
increase the risk of T2DM. At the same time, these DNB genes also participate in many important biological processes related to the T2DM development and progression, such as response to insulin and inositol stimuli, abnormal lipid metabolism and immune system response. From the viewpoint of pathogenic mechanism explanation, DNB genes were also found to play an important role in regulating biological processes. These genes tend to be located at the upstream of pathway so that they may drive the downstream molecules to have the transition change at a transcription level. Besides, some inconspicuous functions identified in the corresponding tissues not only reveal new biological responses stimulated by insulin resistance but also signal the emergence of the critical transitions for T2DM, e.g. steroid hormone biosynthesis is significantly dysregulated in liver and adipose at the critical transition from the normal stage to the pre-disease stage and then final disease-deterioration stage.

Furthermore, the warning signals of disease deterioration detected in different tissues imply some common mechanisms or functional regulations for the similar abnormal situations. For instance, although the dysfunction of steroid hormone biosynthesis does not exist in muscle owing to tissue specificity of this biological function, the dysfunction of hormone response was still detected at the pre-disease stage simultaneously for both liver and adipose. This phenomenon might suggest that tissue-specific DNB analysis is reliable and its result is significance. The studies in this work demonstrate that DNB method not only detects the early warning signals for distinguishing the normal and pre-disease periods during disease development and progression, but also provides the leading network (or the dominant group of genes) of the transitions for revealing molecular mechanisms of disease initiation and progression at a network level. Clearly, the DNB method can be directly applied to the analysis of other complex diseases and also to the study of biological processes in a similar manner [15, 16, 18, 19].

SUPPLEMENTARY DATA
Supplementary data are available online at http://bib.oxfordjournals.org/.

Key Points
- We reviewed recent theoretical results of DNB for early diagnosis of complex diseases.
- We applied the DNB method to the analysis of T2DM.

FUNDING
National Natural Science Foundation of China under nos. 61134013, 61072149, 31200987 and 91029301; Shanghai Pujiang Program; Chief Scientist Program of Shanghai Institutes for Biological Sciences, CAS under no. 2009CSP002; and the FIRST program from JSPS initiated by CSTP.

References


