SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
COMPUTER MODELLING
DATA INTEGRATION
DATA MINING
TYPE 2 DIABETES MELLITUS

Received
8 May 2013

Accepted
28 June 2013

Published
24 July 2013

Correspondence and requests for materials should be addressed to Y.W. (ywang@amss.ac.cn) or L.C. (lnchen@sibs.ac.cn)

Spatio-temporal analysis of type 2 diabetes mellitus based on differential expression networks

Shao-Yan Sun[1,2], Zhi-Ping Liu[1], Tao Zeng[1], Yong Wang[3] & Luonan Chen[1,4]

[1]Key Laboratory of Systems Biology, SIBS-Novo Nordisk Translational Research Centre for PreDiabetes, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China, [2]School of Mathematics and Information, Ludong University, Yantai 264025, China (co-first affiliation), [3]National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, [4]Collaborative Research Center for Innovative Mathematical Modelling, Institute of Industrial Science, University of Tokyo, Tokyo, 153–8505, Japan.

T2DM is complex in its dynamical dependence on multiple tissues, disease states, and factors' interactions. However, most existing work devoted to characterizing its pathophysiology from one static tissue, individual factors, or single state. Here we perform a spatio-temporal analysis on T2DM by developing a new form of molecular network, i.e. 'differential expression network' (DEN), which can reflect phenotype differences at network level. Static DENs show that three tissues (white adipose, skeletal muscle, and liver) all suffer from severe inflammation and perturbed metabolism, among which metabolic functions are seriously affected in liver. Dynamical analysis on DENs reveals metabolic function changes in adipose and liver are consistent with insulin resistance (IR) deterioration. Close investigation on IR pathway identifies 'disease interactions', revealing that IR deterioration is earlier than that on SlC2A4 in adipose and muscle. Our analysis also provides evidence that rising of insulin secretion is the root cause of IR in diabetes.

Type 2 diabetes mellitus (T2DM), which is mainly a glucose metabolism disorder closely associated with modern life style, has become one of the leading health problems in the world[1]. It is widely believed that T2DM is a heterogeneous disease resulting from the complicated interplay of multiple tissues and various factors (genetic, epigenetic, and environmental factors) in a dynamical manner. For instance, one of the main challenges stemming from studying T2DM is to reveal the mechanism of insulin resistance (IR), which is associated with complex interactions of liver, adipose tissue, skeletal muscle, pancreas, kidney, and even brain, and in particular, dynamically affects various biological processes, metabolic networks, and signaling pathways at different stages during the disease progression. However most published works focused on analyzing disease pathophysiology mainly by individual tissues, factors, or states in a static manner, which clearly cannot characterize the essential spatio-temporal features and complex gene-phenotype associations of T2DM[2].

Genes are tightly regulated to execute the proper biological functions in a cell for responding internal or external perturbations[3], and thus their expression variations during disease deterioration process are causally associated with the phenotype changes. Recent rapid advance on high-throughput technologies provides unprecedented opportunities to measure dynamical behaviors of various tissues at a genome-wide level[4,5]. To date, gene expression profiling has been widely used for complex diseases research and a variety of methods have been proposed to identify molecular biomarkers or select features out of the gene expression profiling. Considering the large volume of the data obtained through high-throughput techniques and the negative influence from noise, a natural strategy is to look for those genes that are dramatically different in disease state and normal state. Traditional methods use statistical techniques, such as *t*-test or fold change, to find 'differential genes' (DGs) showing significantly differential expressed patterns between case samples and a control group[6–8]. Though DGs are considered as candidates to play a pathogenic role, they are usually picked out individually, and the output is a ranked list of genes, while dependences or interactions among genes are ignored[3,9].

On the other hand, network-based systems biology offers a quantifiable description of the molecular networks that characterize the complex interactions and the intricate interwoven relationships that govern cellular functions, among those tissues and disease related genes to explain the molecular processes during disease development and progression[10–13]. Network-based approaches have been developed to extract informative genes relying on bio-molecular networks (e.g., PPI network and gene regulatory network), rather than individual genes. The hypothesis underlying these methods is that genes associated with the same disorder tend to share common

functional features, reflected in that their protein products have a tendency to interact with each other. The popular way is based on co-expression relationship to identify the disease-related genes whose expressions are highly correlated through the whole dataset[9,14]. However, genes which show highly correlated patterns (e.g., edges) of expression in one biological state, but not in another, may not be highly correlated across the entire dataset, and therefore would not be picked out by co-expression based methods[15,16]. To overcome this problem, recently, 'differential network' (DN)[4] has been developed to extract disease related edges in a construct-then-compare manner through comparing occurred interactions across different static networks under different conditions.

Different with DG and DN paradigm, here we propose the 'dysfunctional interaction' concept to model underlying disease development and progression. Since genes and gene products function not in isolation but as biochemical or physical interactions, we assume that molecular interactions are perturbed by genetic or epigenetic factors will cause final molecular dysfunctions underlying human disease. Therefore, 'dysfunctional interaction' is defined as the molecular interaction which is significantly changed from wild-type to disease condition. This is very similar to the concept of differentially expressed gene. However a molecular interaction consists of two nodes and one edge, quantitative assessment of its change in different conditions is not straightforward. Here we combine two types of perturbations to define 'dysfunctional interaction' as illustrate in Figure 1 (Part I). Type I dysfunctional interaction is resulting from node perturbations, i.e., differentially expressed nodes will lead to the interaction perturbation. Type II dysfunctional interaction is resulting from edge perturbations, i.e., the co-expression relationship of the molecular interaction is differentially changed. Type II is also referred as 'differential interactions' by directly affect interactions while type I is called 'non-differential interactions' by indirectly changing interactions via node changes. Furthermore, if one dysfunctional interaction directly linked at least one disease genes (reported T2DM related genes), then we define it as 'disease interaction'. Therefore, disease interaction is an interaction who owns close connection with the corresponding disease. Actually, by treating the known disease genes as gold standard, we find that 'differential interactions' and 'non-differential interactions' are complementary in studying T2DM, shown in Figure 1 (Part II).

Given these backgrounds and the dysfunctional interaction concept, this paper aims to conduct an integrative spatio-temporal analysis on cross-tissue interactions of T2DM so as to gain a system-wide understanding for diabetes, by developing a new type of molecular network, i.e., 'differential expression network' (DEN). To construct DEN (Figure 2 A–G), we extract 'differential interactions' and 'non-differential interactions', which are capable of characterizing the initiation and progression of T2DM. In contrast to the traditional 'differential genes' (DG, Figure 2 C) or 'differential network' (DN, Figure 2 D), the DEN is able to better describe phenotype differences at the network level. From the network viewpoint, DEN (Figure 2 G) actually not only covers DG and DN, but also includes disease-related 'non-differential interactions' (Figure 2 E) which are missed in DN. Note that a non-differential interaction is defined as an edge without significant differential strength, but both of its two linked genes have differential expressions between case and control samples, thereby highly associating with the biological process.

To analyze T2DM of Goto-Kakizaki rat (an animal model of T2DM) based on the time-course gene expression profiles (five time points), we first reconstruct the DEN across three classical insulin target tissues, i.e., white adipose, skeletal muscle and liver, at several time points based on the protein-protein interaction (PPI) network combined with gene expression profiles. The three tissues are generally considered as three most important metabolic tissues. As a result, the characterization of T2DM inside each tissue at molecular level is represented by its DEN. Based on these DENs, the commons and differences are identified among these tissues. We are able to demonstrate that DENs provides rich biological insights at gene, interaction, pathway, and even global network structure level.

Results

Proof-of-concept study on dysfunctional interactions. We use animal model of diabetes, Goto-Kakizaki (GK) rat, and its control, Wistar-Kyoto (WKY) rat to study the development and progression

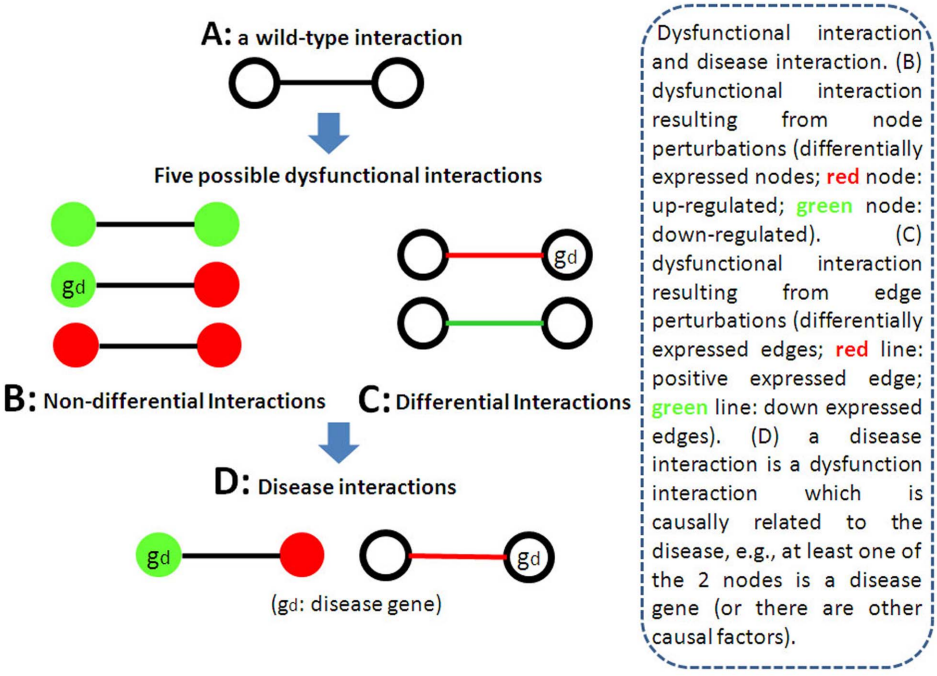


Figure 1 | Dysfunctional interactions and disease interactions. (A) Schematic illustration of Type I disease interaction resulting from node perturbations (differentially expressed nodes, red node: up-regulated, green node: down-regulated). (B) Schematic illustration of Type II disease interaction resulting from edge perturbations (differentially expressed edges, red line: positive expressed edge, green line: down expressed edges).

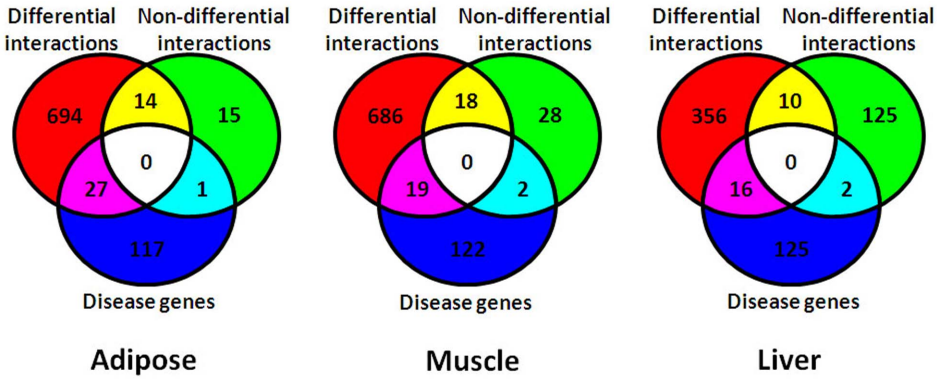


Figure 2 | Venn diagram of three sets of genes from 'non-differential interactions', 'differential interactions', and 'disease genes'.

of T2DM, and the detailed time-course data of gene expressions on three tissues are described in Materials and methods.

The dysfunctional interaction model underlying disease development and progression is shown in Figure 1. As a proof-of-concept example, we firstly identified the 'dysfunctional interactions' of T2DM to demonstrate its usage. Specifically, we picked out all the existing protein-protein interactions. Each interaction was assigned a strength value by the correlation of its linked two genes. Then by comparing the gene expression level in case and control samples, each differential interaction was detected based on its differential strength of the edge, whereas each non-differential interaction was identified based on differential expressions of its two linked genes. Since there is no existing gold-standard for dysfunctional interactions, we validate the identified dysfunctional interactions by checking whether its connecting nodes are known disease genes in Rat Genome databases. In adipose, we predict 56 dysfunctional interactions (1 type I non-differential interactions and 55 type II differential interactions) and they contain 28 know disease genes in all. In muscle, there are 39 dysfunctional interactions (2 type I non-differential interactions and 37 type II differential interactions), and these interactions cover 21 disease genes. While in liver, there are 30 dysfunctional interactions (4 type I non-differential interactions and 26 type II differential interactions) which involve 18 known disease genes. Figure 2 shows the Venn diagram of three sets of genes from 'non-differential interactions', 'differential interactions', and 'disease genes'. It is easy to see that the differential interactions may cover more disease genes, and the covered disease genes due to 'non-differential interaction' and 'differential interaction' appear to be complementary on a certain degree.

As mentioned in introduction and illustrated in Figure 1, if one dysfunctional interaction can cover at least one disease genes (reported T2DM related genes), then it is labeled as 'disease interaction' in our work. Eventually, we obtain 56/39/30 disease interactions for adipose/muscle/liver respectively. All the disease interactions and the covered disease genes (covered_DG) are listed in the Supplementary Table 1 (Supplementary Table 1A for adipose, Supplementary Table 1B for muscle, and Supplementary Table 1C for liver). We then perform literature search for those disease interactions. Indeed, these disease interactions are closely related to the T2DM. The following are some given examples. (1) The interaction between ACP5 and ENPP1 is identified as type 1 dysfunctional interaction for these two genes are both significantly differentially expressed across case and normal conditions in adipose tissue. Indeed, ACP5 has been approved related with diabetes for human and mouse, and remaining to be confirmed for rat. According to our numerical computation, it is significantly down regulated in GK compared with WKY which is consistent with the results published in 2011[17]. ACP5 is associated with macrophages, the lower expression in GK likely means the higher inflammation in GK population. ENPP1, also known as PC-1, is not involved in T2DM related in RGD database released in 2009 used in this paper, but it is considered to be a likely candidate gene for insulin resistance and type 2 diabetes[18]. Therefore, the interaction between ACP5 and ENPP1 is an interaction that show close association with T2DM. Besides, in this tissue, many differential interactions contain gene GSTT1, although this disease gene does not show significant differential expression pattern. While actually, GSTT1 has been reported to play a vital role in the development of T2DM[19]. (2) In skeletal muscle, GSTM1 is overexpressed. This gene is known for its broad range of detoxification and in the metabolism of xenobiotics. It plays a vital role in the development of T2DM[19]. GSTA4 is downregulated under case condition compared to normal state. Though it has not been identified exactly related to T2DM, it may contribute to the development of insulin resistance and type 2 diabetes[20]. Thus, the interaction between GSTM1 and GSTA4 is reasonably considered to be perturbed due to T2DM. Interestingly, the interaction between ACP5 and ENPP1 is also identified in muscle, the different thing is that, here, it is identified for the co-expression of them is significantly changed. The correlation coefficient is 0.8085 under normal state while is reduced to -0.2208 under diabetics, which means that the co-participated function between the corresponding two proteins dysfunctions. (3) In liver, the gene expression of PIK3R1 and NGFR are both significantly down-regulated in GK rats compared to WKY subjects, so the interaction between them is regarded as non-differential interaction. Thus it should own close connection with T2DM. Actually, this viewpoint is reasonable and the supported information is shown as follows. PIK3R1 is a heterodimer containing a regulatory subunit and a catalytic subunit, and has been identified as a T2DM related genes in RGD database. Decreased expression of this gene suggests a possible decrease in insulin signaling at the level of PI3-kinase. NGFR is a nerve growth factor receptor. Its decrease in the genetic expression might be responsible for the pathogenesis of diabetic cystopathy, while hyperglycaemia is part of the cause of these changes[21]. The interaction between GSTM1 and GSTA3 is shown here as another example. This interaction is identified as type 2 dysfunctional interaction since the correlated coefficient is changed from 0.3200 (control) to 0.7231 (case). Among the linked genes, GSTM1 has been approved a T2DM related gene, and as reported in the previous study[22], GSTA3 may be a contributing factor to the increased oxidative stress and incidence of hepatic disease that may occur during diabetes. In addition, it should be noted that the interaction between GSTM1 and GSTA4 is detected as dysfunctional interaction in skeletal muscle, while in liver, the interaction between GSTM1 and GSTA3 is identified. This reflects the common and specific among the two tissues to some degree.

Constructing differential expression network (DEN). We have constructed a tissue-dependent 'differential expression network'

(DEN) by complementarily considering both individually genes and dysfunctional gene pairs. The detailed procedure to construct DEN is illustrated in Figure 3, where the comparison of DEN with traditional schemes is also shown. A DEN is composed of 'non-differential interactions' (Figure 3 E) and 'differential interactions' (Figure 3 F), which cover both 'differential genes' (DG, Figure 3 C) and 'differential network' (DN, Figure 3 D). In particular, a DEN includes the non-differential interactions which are missed in the DN.

Specifically, by combining a molecular network (e.g. a PPI network) and gene expression profiles, we first constructed the molecular network under case or control condition, where the strength of an edge can be given by the correlation of its linked two genes. Then by comparing case and control samples, each differential interaction was detected based on its differential strength of the edge, whereas each non-differential interaction was identified based on differential expressions of its two linked genes. Note that a non-differential interaction is not a differential edge, but both of its two linked genes have differential expressions, thereby relating to the disease process. We will show that our scheme outperforms both 'differential genes' and 'differential network' schemes in T2DM study.

Extracting differential genes and non-differential interactions. We adopted a commonly employed technique to identify significant differentially expressed genes noted as 'differential genes' by comparing case and control samples, i.e., GK and WKY rats in our study. We identified 700, 751, and 907 significantly differentially expressed genes (with multiple testing adjustment cutoff 0.05), for adipose, muscle, and liver respectively, covering 6, 4 and 8 out of the 143 known disease genes (literature reported T2DM related genes). Furthermore, we filtered the unreliable differential expressed genes through discarding some trashy differential genes by protein interaction information, since protein-protein network was steadily constructed experimentally characterizing the physical interactions between proteins. Specifically, we mapped the differential genes from different tissues to PPI network respectively. Only the edges whose two nodes are all differentially expressed were further considered and denoted as 'non-differential interactions', which are not incorporated in the DN. Consequently, we obtained 46/55/138 'non-differential interactions' among 30/48/70 proteins for each tissue (adipose/muscle/liver), covering 1/2/2 disease genes. It is clearly that the coverage for disease genes is raised due to the incorporation of PPI network information.

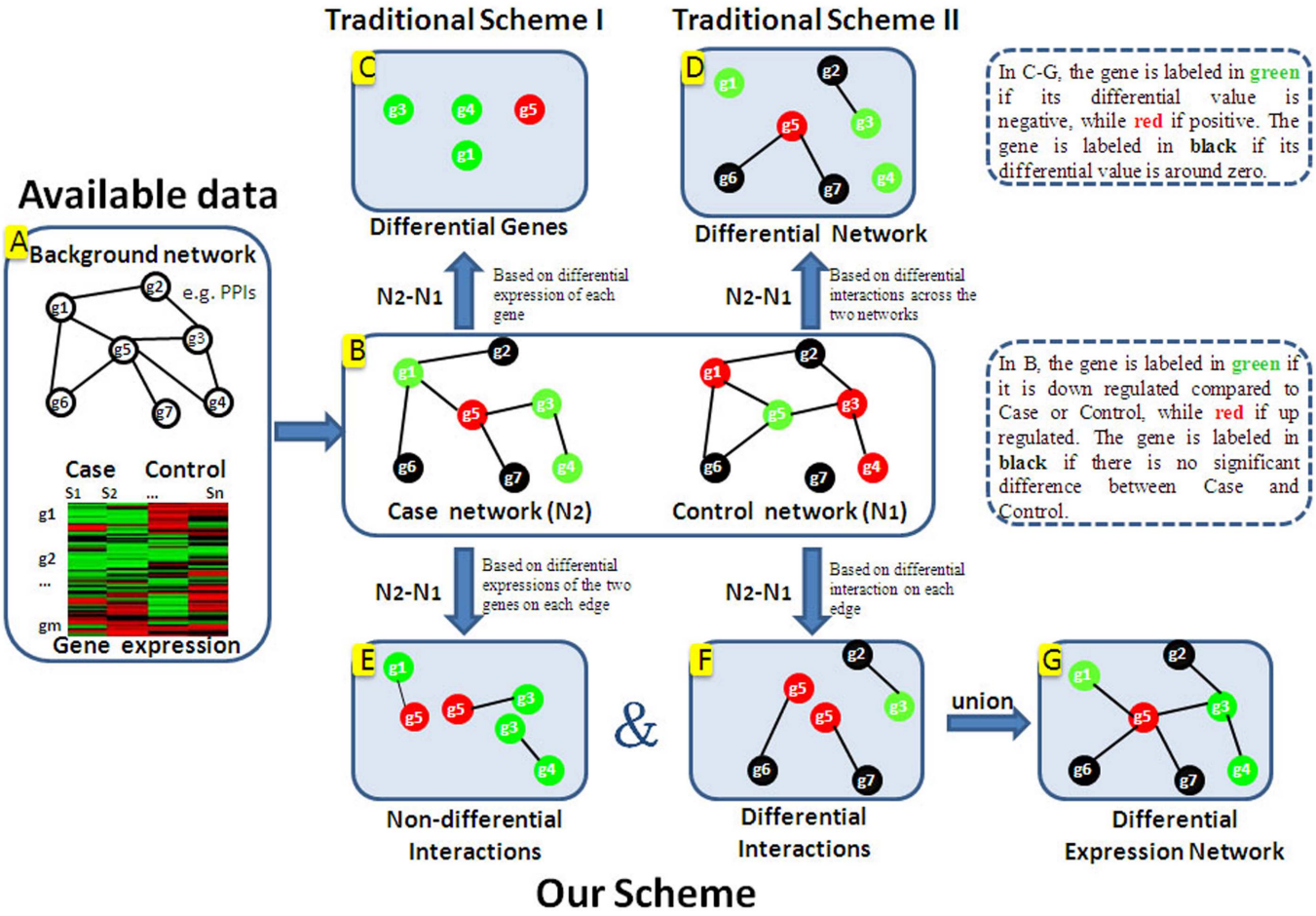
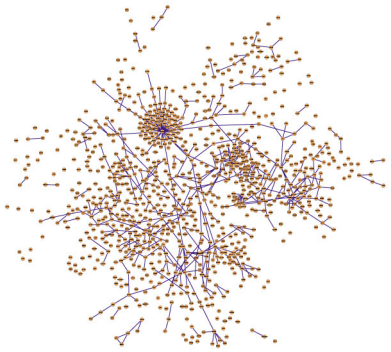


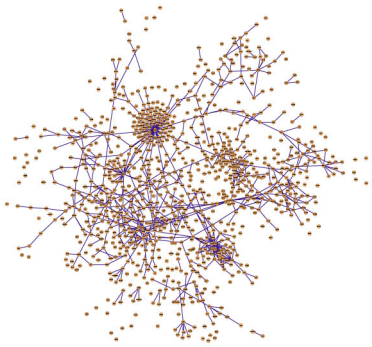
Figure 3 | Scheme of constructing 'differential expression network (DEN)'. The new form of network named DEN is based on both molecular network (e.g. PPI network) and gene expression profiles to characterize biological process (Fig. 2A), in contrast to traditional 'differential genes' (DG; Fig. 2C) or 'differential network' (DN; Fig. 2D). DEN is composed of 'non-differential interactions' and 'differential interactions'. In Fig. 2 B, the gene is labeled in green if it is down regulated compared to Case or Control, while red if up regulated. The gene is labeled in black if there is no significant difference between Case and Control. In Fig. 2 C–G, the gene is labeled in green if its differential value is negative, while red if positive. The gene is labeled in black if its differential value is around zero. As a result, DEN is constructed for each tissue. In addition, the comparison of our DEN scheme with traditional schemes is also can be shown. DG only considers those individual genes whose expressions have significant differences under two different conditions, while DN is composed of the differential interactions that have significant differences in different states, without consideration of disease-related non-differential interactions.

A

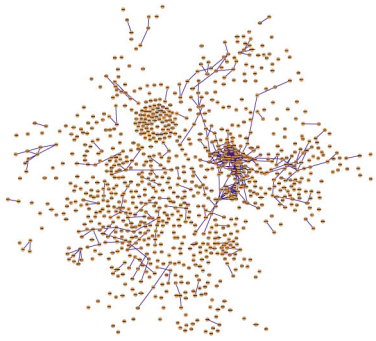
Adipose



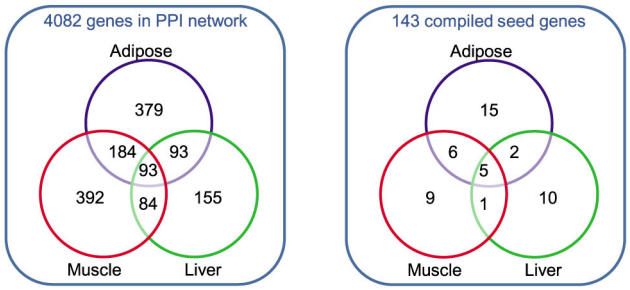
Muscle



Liver



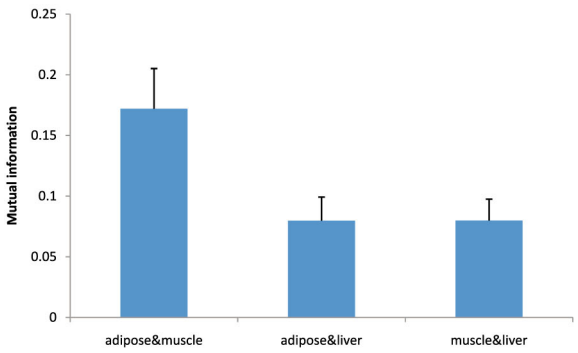
B **Venn diagrams for node sets from three DENs at two different resolutions**



C **Comparison of three DENs based on network topology**

Items	adipose		muscle		liver	
	full	*comp*	full	*comp*	full	*comp*
Nodes	749	286	753	464	425	105
<*k*>	1.768	2.210	2.109	2.5	2.207	4.229
<*l*>	10.553	10.762	7.789	7.821	4.606	4.860
<*c*>	0.024	0.022	0.031	0.030	0.079	0.162

D **Similarity of three tissues based on the relative expressions for the common genes**



E **Functional analysis based on enrichment pathways for three tissues**

Items	Adipose	Muscle	Liver
No. of pathways in metabolism **(percentage)**	16 **(30.19%)**	8 **(15.38%)**	21 **(61.76%)**
No. of pathways in immune **(percentage)**	8 **(15.09%)**	6 **(11.54%)**	5 **(14.71%)**
No. of pathways in cancer **(percentage)**	11 **(20.75%)**	13 **(25.00%)**	3 **(8.82%)**
No. of pathways enriched in total	53	52	34

Figure 4 | Systems analysis for tissue specific DENs. (A) Showing the largest component of DEN for each tissue. (B)Venn diagrams show the overlap of genes perturbed in three tissues. (C) Main topological parameters for each tissue's DEN and corresponding giant connected component. (D) Similarity between any two tissues measured by using mutual information. (E) Percentage of enriched pathways in three major functional classes.

Extracting differential interactions. Instead of choosing genes individually, here we chose genes in a pairwise way, i.e., interactions. An interaction (corresponding to an edge in a network) will be picked out if its two directly linked genes are strongly correlated (the Spearman correlation coefficient is greater than or equal to some threshold, such as 0.7) under one strain while not (e.g. less than 0.5) in the other condition. The difference embodied in co-expression implicates that the interaction between this gene pair is perturbed due to the disease, and such an interaction is named as 'differential interaction'. Through screening all edges in PPI network, we have identified 623/741/340 differential interactions among 733/723/382 proteins for adipose/muscle/liver respectively, containing 27/19/16 disease genes.

Interestingly, there is no overlap between the two sets of covered disease genes picked out by our 'non-differential interactions' and 'differential interactions' respectively. This fact indicates that the two type interactions are complementary (see the Venn diagrams in Figure 2).

Integrating differential and non-differential interactions. Since the aforementioned two types of interactions to identify disease related gene candidates are complementary, here, by combining them together we propose a new form of molecular network, 'differential expression network' (DEN), which is capable of characterizing each tissue's alterations between case and normal samples. Specifically, the DEN is built by taking the union edges of non-differential interactions and differential interactions.

For each tissue, we obtained a differential expression network composed of 749 (662), 753 (794), and 425 (469) nodes (edges), respectively, which contained 28/21/18 (adipose/muscle/liver) disease genes. The high coverage for disease genes confirms the hypothesis that more accurate disease related information can be captured by our DENs. Figure 4A displays the largest component of three tissue-dependent DENs.

Static analysis on tissue-dependent DENs. In the following, we performed systematic analyses on these DENs from global to local level.

Macroscopic topology analysis. Firstly, we studied the main topological measurements of each tissue's full DEN (denoted as 'full') and its giant connected component (denoted as 'comp'). The main topological measurements are average degree (<k>), average length of shortest paths (<l>), and average clustering coefficient (<c>). The detailed information is shown in Figure 4C, where the numbers of nodes are all listed to provide a reference for the topological parameters from different tissue's DENs.

From the global picture of network and the topological measures, we conclude that muscle and adipose present a consistent tendency, while liver shows distinct characters. Liver's DEN has a higher average degree <k> than those from other two tissues, which is especially notable when we calculate <k> on giant connected components. Again on average, the <l> of DENs from adipose and muscle are comparable, while the liver's DEN displays the lowest value for <l>. This fact indicates that two proteins in the DEN of liver tend to be closely connected. In addition, higher average clustering coefficient (<c>) was observed in DEN of liver, which indicates a stronger tendency of proteins in liver to form tight clusters or groups. All the topological observations suggest that there are no big differences in the macroscopic topological properties among the DENs from adipose and muscle, indicating the certain homogeneity between these two tissues during the disease progression.

Gene level analysis. To find possible relationships among the three tissues under T2DM, we simply counted the number of corresponding disease gene sets inside each tissue and checked their pairwise overlaps. The Venn diagrams are shown in Figure 4B. It shows that every two tissues have significantly overlapped genes (p-value $<$ 1e-32) in their DENs. This demonstrates that three tissues are tightly associated by sharing common T2DM phenotype. However, we found that the overlap between adipose and muscle is significant (p-value $<$ 0.01) when we refined all the identified genes into 143 disease genes, which can represent the biological processes in T2DM more accurately.

Furthermore, we compared three tissues more concretely by investigating the changes in gene expression patterns. For this purpose, we introduced a notion named relative expression which can display the changes of gene expression from control condition to disease condition. Then we calculated the mutual information (MI) of the same gene between different tissues. In this way we quantitatively measured the similarity of tissues on the relative expression pattern level for one specific gene. Accumulating the mutual information of 93 common genes allows us to measure the overall similarity of pairwise tissues. The mean and standard error for the MIs of 93 common genes from three tissues are shown in Figure 4D. Consistent with previous findings, the most similar patterns of expression are observed between adipose and muscle, as shown in Figure 4D.

Functional analysis. Next, we performed functional analysis on the DENs underlying T2DM. The pathways enriched in each tissue's DEN were picked out by the online software DAVID. Then we compared the pathway enrichment results among different tissues.

As a result, we found 53, 52, and 34 significantly enriched pathways (Bonferroni-adjusted p-value $<$ 0.05) in adipose, muscle, and liver respectively. All these pathways were manually classified according to the KEGG pathway database, and the percentages of pathways that fell into the same (or different) categories were counted. As shown in Figure 4E, the detected pathways mainly belong to three major classes, i.e., metabolism, immune, and cancer. The percentages of the enriched pathways in metabolism class are 30.19%, 15.38%, and 61.76% for adipose, muscle, and liver. This indicates that the metabolic related functions of liver are disturbed most severely due to T2DM. On the other hand, the percentages of the enriched pathways in immune class of the three tissues are comparable, which implies that the severities of inflammation displayed in these tissues are almost at the identical level.

Gene expression profile analysis. Unsupervised hierarchical clustering of the gene expression data across the three tissues reveals interesting patterns in the gene expression heat map (Figure 5A). First, the adipose, muscle, and liver are clearly distinguishable. Second, the adipose is more close to muscle. Thirdly, the gene expression profiles of GKs and WKYs are closely grouped together for each tissue, while the expression profiles from different disease states are clustered more closely with each other than with the profiles from the corresponding GK and WKY.

The main sources of variation in the patterns of gene expression samples can be visualized as 2D maps (Figure 5B) using principal components analysis (PCA), in which each dimension represents a group of coordinately regulated genes. Principal components (PCs) are a mathematical abstraction. In brief, given a matrix of gene expression, the first PC is the set of genes that accounts for most of the variation among the samples. The second PC is the next set of gene that account for most of the remaining variation, in a manner that is independent of the first PC. Additional PCs account for further variation. In this way, PCA can mathematically identify the main sources of variation in a complex matrix. Figure 5B nicely shows that the first two PCs appear to have meaningful biological associations. Principal component #1 represents the tissue specificity of the samples and the principal component #2 indicates the T2DM state. Using this analysis, samples from one tissue are clearly separated with those from another tissue, based on the tissue specific genes that make up the first PC (PC1). Intriguingly, PC2 divides

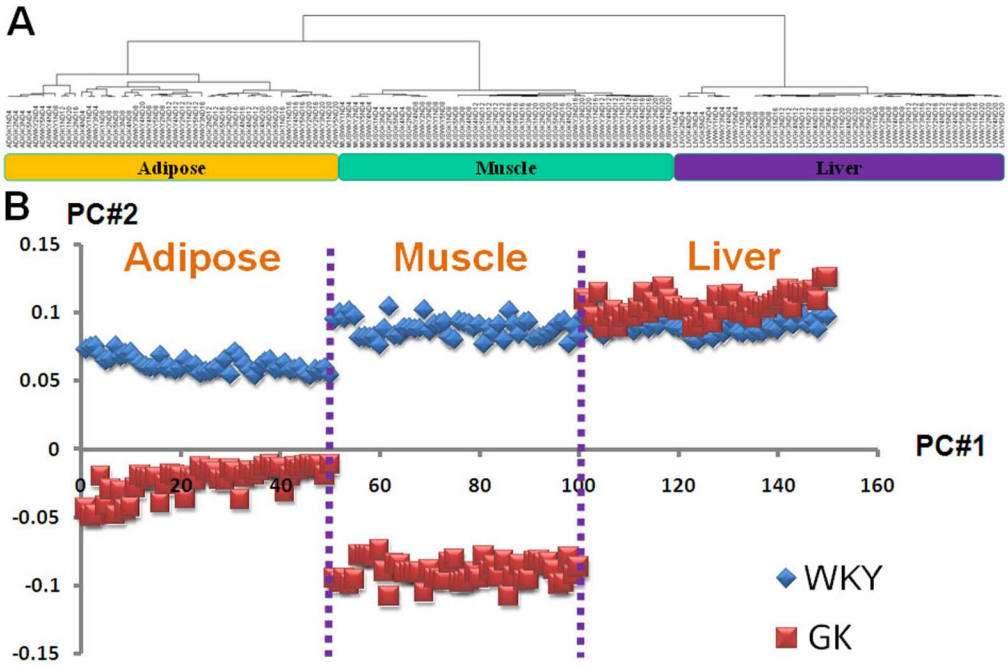


Figure 5 | Global analysis of the gene expression profiles. (A) Unsupervised hierarchical clustering of the gene expression data across the three tissues reveals patterns in the gene expression. (B) The main sources of variation in the patterns of gene expression samples can be visualized as plot using principal components analysis (PCA), in which each dimension represents a group of coordinately regulated genes.

the samples into T2DM and normal classes according to the disease state. The first class consists of cells from WKY samples, and the second class contains GK samples. Basically, there are two main sources of variance in the gene expression data. One is the tissue specificity and another one is the disease state. This highlights the motivation of this study to consider tissue specificity in T2DM study, and gene expression analysis also supports our previous finding that liver is different from adipose and muscle during the progression of T2DM.

Local analysis on insulin resistance pathway. In addition to the global characterization, we also checked the similarities and differences among three key tissues at local pathway level. We focused on the insulin resistance (IR) pathway (from KEGG), which is known as a major player in adipose tissue, skeletal muscle, and liver for diabetes subjects[23]. Although the molecular mechanism of IR pathway has been firmly established, it is unclear if there are tissue-specific differences[24]. For example, though it has been recognized that insulin receptor substrates (IRSs, including IRS1, IRS2, IRS3, and IRS4) are a family of adaptor proteins that are crucial for the action of insulin, it is still not clear whether or not different IRSs play redundant or selective roles in insulin action in different tissues[25]. Our spatial analysis may provide clues for the important interactions leading to insulin resistance.

We characterized tissue's states under two contrast conditions by labeling differentially expressed genes and rewriting interactions. Specifically, one gene is labeled in green, if it is significantly down-regulated in insulin resistant state (corresponding to the subfigures at the bottom in Figure 6). Alternatively, a gene is labeled in red if it is up-regulated. Similarly, a green edge means that the corresponding two genes are strongly and negatively correlated, while red indicates strongly and positively correlated. The results for three tissues are shown in Figure 6.

Again, similarities were observed in adipose tissue and skeletal muscle. Firstly, more changes were detected with respect to node IRS1 in both adipose and muscle tissues from the top subfigures in Figure 6. This indicates the fact that IRS1 contributes to mediating the effect of insulin on glucose transport[26,27]. Consequently, the functions of SLC2A4 have been disordered in these two tissues, while not in liver. As reported, this gene encodes the insulin-regulated glucose transporter (GLUT4) found in adipose tissue and skeletal muscle, which is responsible for insulin-regulated glucose translocation into the cell[28].

More tissue specific alterations were observed by a closer check. (1) We can see from the two pathways in adipose and muscle tissue under normal condition that IRS1 is the main docking protein for the binding and activation of phosphatidylinositol 3-kinases (PI3K). While in GK samples, IRS2 was significantly up-regulated and became the main docking protein which requires a higher insulin concentration than that is needed for a similar binding to IRS1[29]. This confirms that IRS1 and IRS2 are structurally similar and have complementary functions[30]. In addition, we found that the functions of IRS3 were also influenced, possibly due to its corresponding protein to activate the metabolic action of insulin and to promote translocation of GLUT4[31]. (2) IRS1 appears to have its major role in skeletal muscle which corresponds to the viewpoints based on the mice model[32]. As shown in subfigures B1-B2 in Figure 6, a disturbed interaction appears between insulin receptor (INSR) and IRS1, and the IRS1-associated PIK3R1 is significantly decreased (indeed, PIK3R1 is observed significantly down-regulated across three tissues). In addition, the insulin signaling at the level of IRS1/PIK3R1 is disrupted from GK rats compared with that of WKY animals, which has been detected in mice models with T2DM[33]. (3) In liver, only IRS3 seems to have significant alterations. Previous studies have reported that in contrast to IRS1 and IRS2, IRS3 is a smaller molecule and has fewer phosphorylation sites[30]. Therefore, IRS3 might function differently from IRS1 and IRS2, which are capable of mediating insulin's action to promote translocation of GLUT4 glucose transporters. Since GLUT4 is the insulin-regulated glucose transporter found in adipose tissue and skeletal muscle while not in liver, it is reasonable to consider that IRS3 also undertakes other distinctive functions from IRS1 and IRS2 in GK liver.

Dynamical analysis on tissue-dependent DENs. Furthermore, we made a pilot investigation from the spatio-temporal (that is, tissue

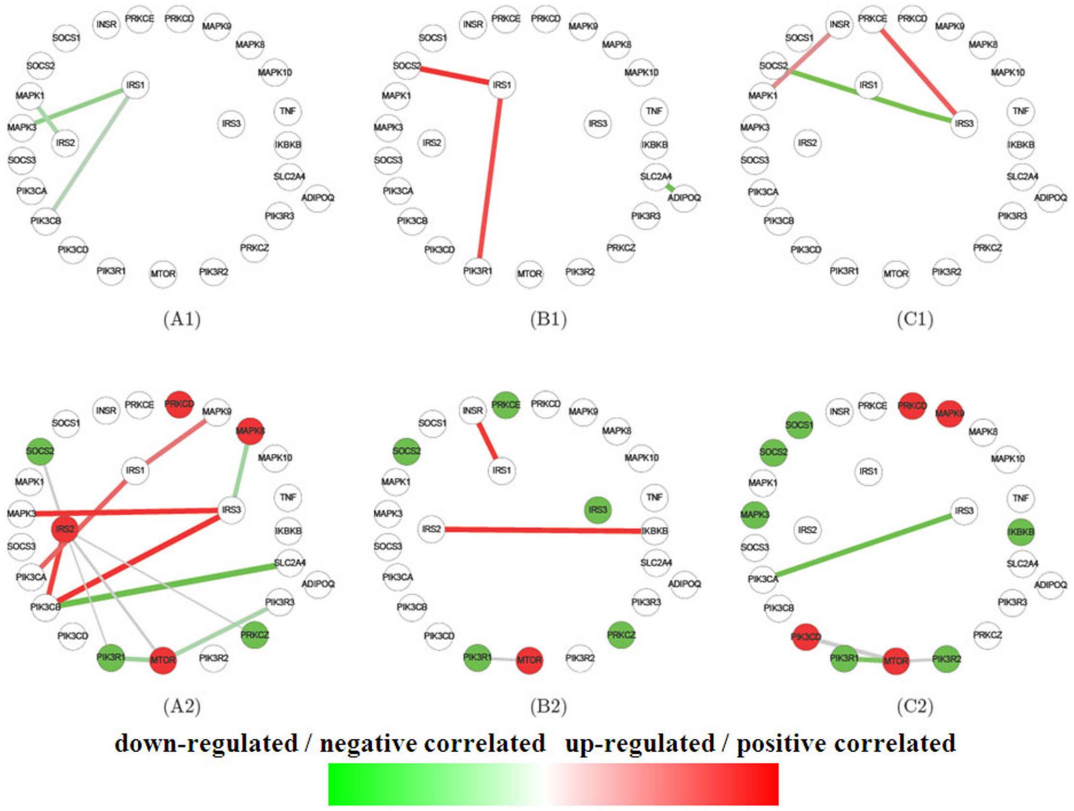


Figure 6 | Different states of insulin resistant pathway under two contrast conditions in three tissues. From left to right, the subnetworks correspond to adipose, muscle, and liver respectively. The subnetworks in the top level correspond to normal condition (WKY), and the subnetworks in disease condition (GK) are listed at the bottom level. Genes from WKY tissues are colored in white, while in green/red according to down/up-regulated sitations in GK. Given one tissue, all the differential interactions in IR pathway are casted to normal or disease strain. A green edge means that the corresponding two genes are strongly and negatively correlated, while red one indicates strongly positive correlation. The gray edges mean that though they are not differential interactions, but their two nodes (genes) are simultaneously differentially expressed.

and age) perspective to study and compare the dynamic changes occurred in three tissues during the development of T2DM.

Temporal analysis on global DENs. For each tissue, we constructed one differential expression network (DEN) at each time point. As a result, 5 DENs in different periods, which span the stages from 4 wk (i.e., week) to 20 wk, are created with the interval 4 weeks. In other words, we have constructed 5 time-course differential expression networks for each tissue to characterize the dynamical changes. These networks are spatio-temporal dependent since they are based on the gene expressions localized on corresponding stage from corresponding tissue.

We firstly analyzed the dynamic topological properties of these time-course DENs. The results are summarized in Figure 7. It shows that the size of time-course differential networks from the same tissue was relative steadily except for three singular points (adipose at the last stage, muscle and liver at the first stage (Figure 6A)). Further observations from Figure 7B–D indicate that, (1) for adipose, both the average degree and clustering coefficient increase from 4 wk to 12 wk and then decrease; while the shortest path shows a peak value at 4 wk, then decline in the next stage and keep steadily in the following stages. (2) For muscle, both the average degree and the clustering coefficient subject to a parabola with time, and touch the bottom at 12 wk or 16 wk. The shortest path varies slightly from 8 wk to 20 wk. (3) For liver, the average degree, the clustering coefficient, and the shortest path all fluctuate, specifically, three parameters increase from 8 wk to 12 wk, then decrease from 12 wk to 16 wk and again increase from 16 wk to 20 wk. These results suggest that the dynamic changes underlying topological properties of time-course DENs are similar to a certain degree in adipose and muscle.

Next we investigate the dynamical changes of the biological functions underlying T2DM. The percentages of enriched pathways were taken as the index to characterize functions underlying each differential networks. Genes in the 5 time-course differential networks were used to reveal enriched KEGG pathways. Dynamic functional indexes regarding to metabolism and immune are shown in Figures 7E and Figure 7F. We found that the dynamic changes upon metabolism and immune seem to be complementary in adipose, also in liver. In these two tissues, the dynamic metabolic index increases from 4 wk to 12 wk and then decreases by reaching its peak value at 12 wk. Interestingly, this change tendency is consistent with the dynamic changes of HOMA-IR index (one index that is generally accepted as a measure of whole body insulin resistance) examined by Almon group[34]. This indicates a positive correlation between the insulin resistance and affected metabolic functions in adipose and liver. However, an unexpected arise occurs in liver at 20 wk. It might be due to the increased size of differential network at that time. In addition, muscle shows its minimum in metabolism index while its maximum in immune index at 12 wk, which is different from the other two tissues.

Temporal analysis on IR pathway. Again we focused on the IR pathway and screened every existed interaction in IR pathway at each stage for three tissues. We aim to check if the interaction is disturbed at some stage in each tissue. To reduce the impact from noise, we only counted those disturbed interactions whose dynamic changes are stable in two or more continuous time points. Similar to

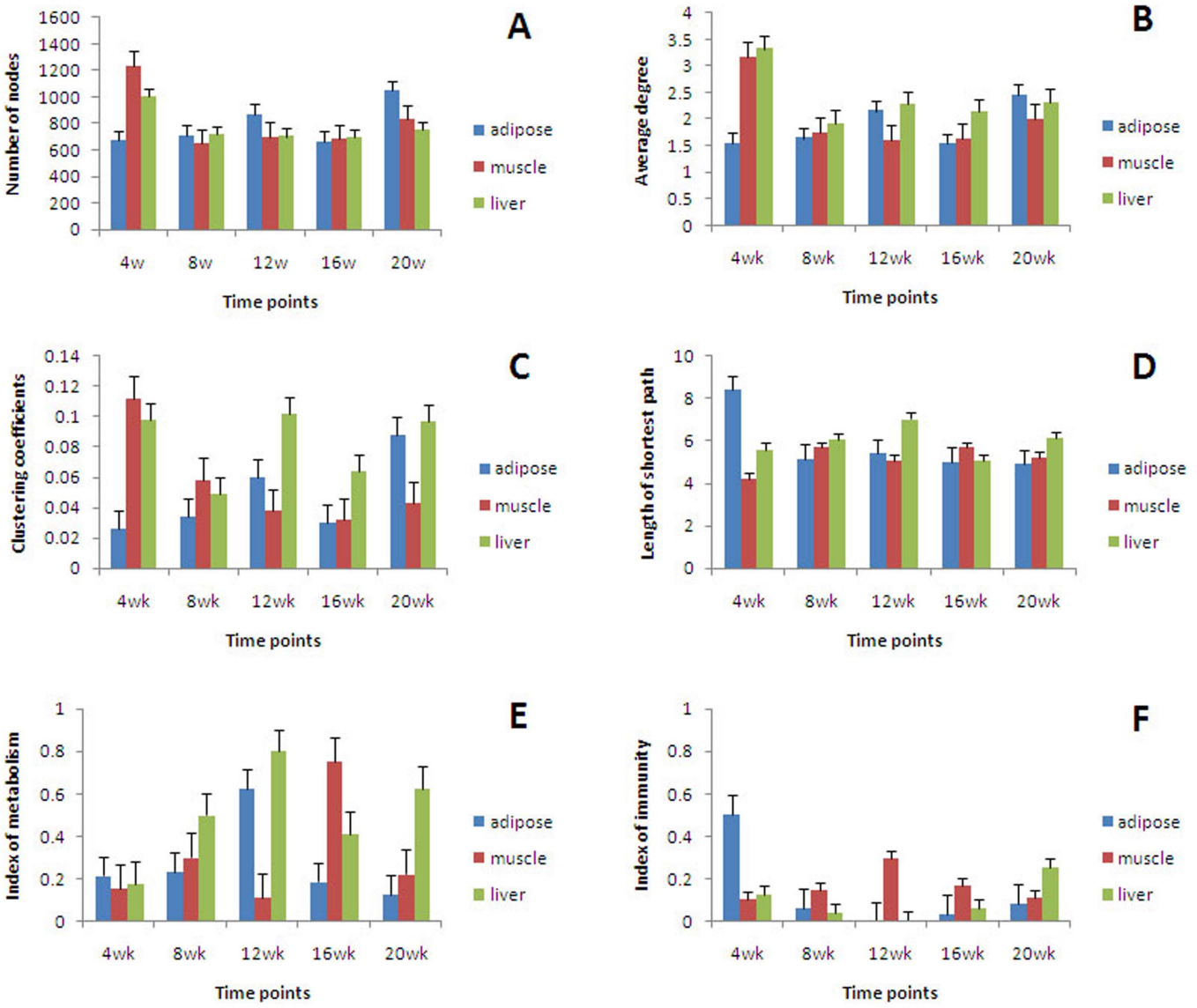


Figure 7 | Topological properties of time-course DENs for three tissues. (A) Number of nodes of the corresponding DENs. (B) Average degree of the corresponding DENs. (C) Clustering coefficient of the DENs. (D) Length of the shortest path of the DENs. (E) Percentages in metabolic function class enriched in DENs. (F) Percentages of the enriched pathways in immune function class enriched in DENs.

the definition in static DEN, such interactions are defined as 'dysfunctional interactions', which are considered to be a stable form to distinguish disease and normal samples in contrast to the traditional 'disease genes'. Here, 'stable change' means that the direction (up or down) of regulating is consistent in different time points, regardless of the absolute value of gene expression level for single gene and correlation coefficient for gene pair. We list the dysfunctional interactions and corresponding time points in Figure 8 and label them using different colors corresponding to different tissues. We found some interesting facts: (1) In adipose, the abnormal perturbations related to IRS1 and IRS3 occur in the earlier stages, compared with perturbations related to IRS2. (2) The perturbations associated with SLC2A4 appear in the later stages (from 12 weeks on), as well in skeletal muscle. (3) The interaction between IRSs and PI3K was disrupted as shown in IR pathway map from KEGG. In adipose, the disruption between IRS2 and PIK3R1 occurred from 16 weeks to 20 weeks. In muscle, the interaction between IRS1 and PIK3R1 was disrupted through all the 5 points in time. In liver, the interaction was disordered at the first 2 time points.

In addition, our temporal analysis supports that rising of insulin secretion is the root cause of insulin resistance in diabetes. It has been a long standing problem in diabetes study for the relationships between insulin secretion and insulin resistance[41]. It's well-known for the strong relationship between basal insulin levels and diabetes. Increasing fasting insulin levels have been documented as factor to impaired glucose tolerance and severe diabetes. However, this correlation provides no information on causation. Our temporal analysis in a tissue-specific manner may provide some hints for this important problem. We check the insulin receptor substrate 1 (Irs1) to represent insulin resistance, since it plays a key role in transmitting signals from the insulin and insulin-like growth factor-1 (IGF-1) receptors to intracellular pathways PI3K/Akt and Erk MAP kinase pathways[42]. For insulin secretion, transcription factor 7-like 2, also known as TCF4, is chosen since its critical role in insulin secretion and linkage to higher risk to develop type 2 diabetes[43]. As shown in Figure 9, our data support that hypersecretion of insulin can precede and cause insulin resistance. Specifically, the peak of insulin resistance (characterized by Irs1) occurring at D16 lags behind the peak of insulin secretion (characterized by Tcf4)

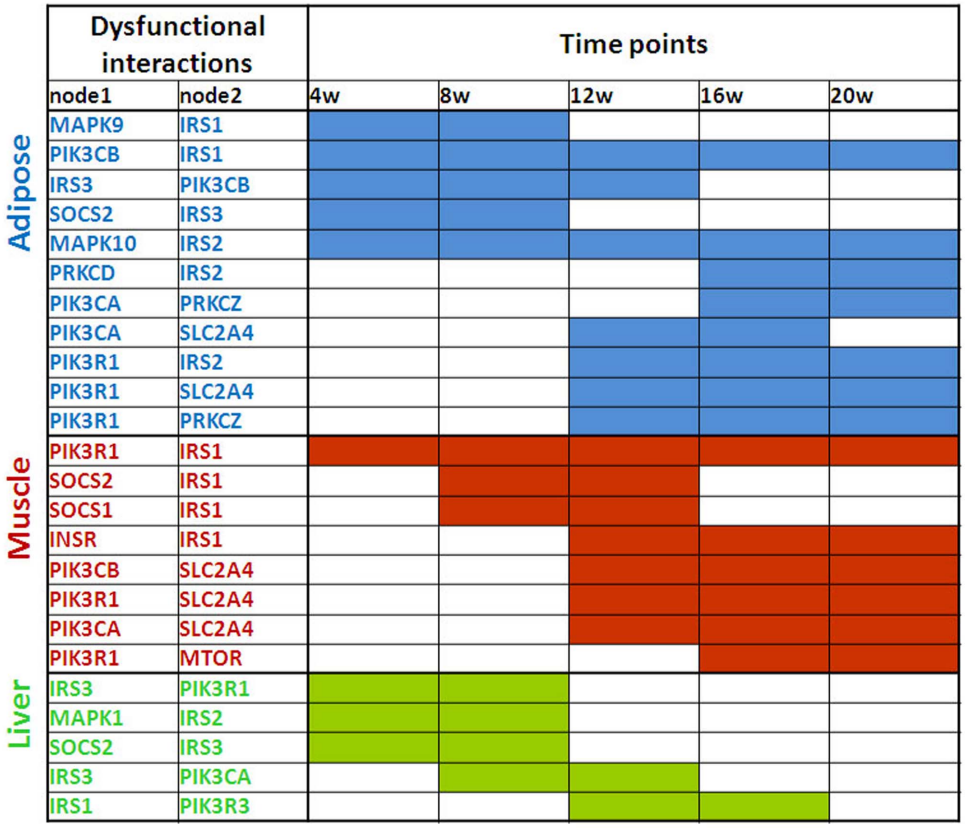


Figure 8 | **Temporal dysfunctional interactions and their corresponding time points in IR pathway in three tissues.**

occurring at D12. This fact supports the model of b-cell secretion of insulin leading to hyperinsulinemia and causing insulin resistance[41].

Discussion

In this paper, we developed a new form of molecular network to characterize the phenotype differences of samples as well as interaction rewiring, and provided not only evidence for the tissue specificity at the molecular level, but also biological insights into the development and progression of T2DM at the network level. Specifically, we present a macroscopic relationship among the three key tissues during diabetes progression at molecular level. Various evidence, i.e., differential expression network, gene overlap, functional analysis, gene expression profile, and IR pathway, all indicates that these three tissues are involved in T2DM progression but play their roles in different ways. Adipose and muscle are two similar tissues in terms of those features, which are slightly different from liver. The dynamical analysis reveals that adipose dysfunctions at an early stage, while liver and muscle dysfunction in all periods.

Type 2 diabetes has long been viewed as a metabolic disease linked to obesity and a sedentary lifestyle. We believe that the traditional disease-causing mutations based model is limited in fully reconcile with the increasingly appreciated prevalence of complex genotype-to-phenotype associations. Here we start from the biomolecular network perturbation model and uncover the differentially changed molecular interactions. Our DEN model cover the edgetic perturbation[35], which mainly cares about the distinct network perturbations by mutation resulting from complete loss of gene products (node removal) or specific alterations in distinct molecular interaction(s) (edgetic perturbation). While our DEN model is general to any driving forces to rewire molecular interactions including genetic and epigenetic factors. Our results demonstrate its advantages in studying T2DM.

In the conventional method, the genes are used to characterize the cells, called gene expression signatures. They are first clustered into some groups based on their expression values, and then each gene list of the groups is characterized by the previous knowledge of biological functions, such as the GO terms. Finally, the terms of the biological functions are interpreted to consider the cellular relationship. In contrast, we developed a new framework by constructing 'differential expression network' (DEN) to investigate the functional relationships between distinctive tissues and further to characterize the spatial-temporal differences under T2DM at network level. Based on these constructed DENs, we firstly performed a systematic comparison analysis across several key target tissues. The results based on static DENs suggest a higher similarity between adipose tissue and skeletal muscle at the level of gene expression and network topology. Such results indicate that the molecular mechanisms underlying T2DM of these two tissues are similar to some extent. In addition, we found that metabolic related functions were affected severely in liver, moderately in adipose and weakly in muscle. Moreover, we also detected that these three tissues were subjected to serious inflammation at nearly identical level. Focusing on the insulin resistance pathway, we found that some proteins belonging to the same family show tissue-specific characteristics, such as IRS1, IRS2, IRS3, etc. Furthermore, analysis based on dynamic DENs implicates many interesting biological phenomena. At the stage when GK rats show most serious insulin resistance, our measured metabolic functions related to T2DM also reach a peak in adipose and liver. Interestingly, the dynamic changes are consistent between the insulin resistance and the affected metabolic function in these two tissues. Observing the dynamic changes in insulin resistance pathway, we found that in adipose and muscle, the functions linked to IRSs (insulin receptor substrates) seem to be affected earlier than those linked to SLC2A4 (a gene encodes a protein that functions as an insulin-regulated facilitative glucose transporter). All these results enable us to further

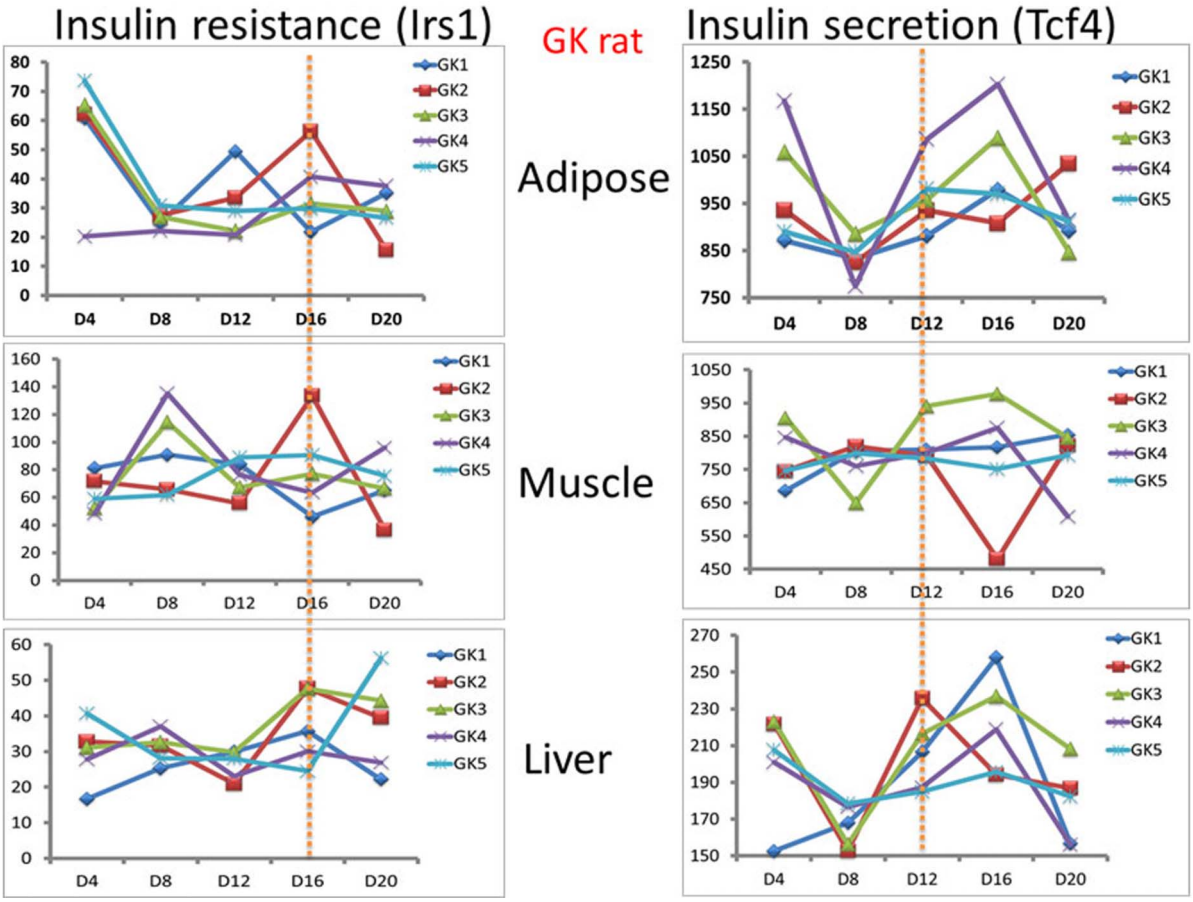


Figure 9 | Our temporal analysis supports that rising of insulin secretion is the root cause of insulin resistance in diabetes. The peak of insulin resistance (characterized by Irs1) occurring at D16 lags behind the peak of insulin secretion occurring at D12. This fact supports the model of b-cell secretion of insulin leading to hyperinsulinemia and causing insulin resistance[41].

experimentally understand the disease, and we expect the emergence of tissue-specific and isoform-specific gene (specifically genes from IRs) knockout studies to corroborate our conclusions.

It should be noticed that our DEN scheme provides a novel way to predict disease genes and further disease interactions. We already demonstrated that this scheme can cover 3–4 folds more known disease genes than traditional differential expressed genes. The research on this area is expected to further demonstrate the effectiveness of DEN framework and is in progress[36,37]. One step further, DEN can be directly applied to the identification of network biomarkers for disease prognosis. Our DEN framework is similar to the network biomarker identification procedure proposed in[38] by integrating PPI and gene expression data. However, the constructed network in[38] only includes a group of genes and the connection pattern among these genes, and co-expression changes in protein interactions, i.e., non-differential interactions are not used in disease classification. In contrast, our differential expression network fully explores all disease-related interactions including non-differential interactions, among the proteins and therefore is able to identify disease genes and disease interactions in an accurate manner.

Methods

Collecting data. We obtained gene expression profiles of three rat tissues (white adipose, skeletal muscle, and liver) from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus database (GSE 13271)[17,34,39]. The gene expression data was measured from normal (WKY) and diseased rats (GK) aged from 4 weeks (4 wk) to 20 wk, and the time interval was 4 weeks. The microarray was composed of 31,099 probes. Our first filter eliminated the probe sets without corresponding official symbol, leaving 25,345 genes for further consideration.

We downloaded the T2DM related genes (referred to as 'disease genes' for convenience in this paper) from the Rat Genome Database (http://rgd.mcw.edu/wg/home). Only those genes whose expressions have been measured in GSE17271 were considered in our analysis. In total, 143 disease genes were analyzed.

The rat protein-protein interaction network (rat PPI) used in this study was integrated from KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway and BioGrid (Biological General Repository for Interaction Datasets). The integrated PPI network contains 24,503 edges (interactions) among 4,082 nodes (proteins).

Identifying differentially expressed genes. It is well confirmed that the propensity of many diseases can be reflected in a difference of gene expression levels in particular cell types[40]. For this reason, genes showing a different expression levels in control crowds (i.e. WKY rats in this paper) and case strains (i.e. GK rats in this paper) are likely related to the disease.

Here the student's *t*-test was utilized to identify differentially expressed genes or simply differential genes. For normal or abnormal group, gene expressions from different samples were taken as replicated experiments. For a gene, whether there is a significant difference in log2 expression level from GK versus WKY can be estimated by *t*-test. The logarithm transformation is to make the data approximately subjected to normal distribution. To get a more powerful and less subject to bias, multiple testing was employed in this work. The threshold for adjusted p-value was set to 0.05.

Identifying differential and non-differential interactions. In addition to disease related genes by screening genes individually, we further detected candidate genes and interactions from the network perspective since genes usually do not function in isolation. Genes associated with the same disorder tend to share common functional features, reflected in the fact that their protein products have a tendency to interact with each other. If some genes have close connections with one disease, it is reasonable to assume that perturbations would invade into the interactions among them. We identified such interactions based on 'differential interactions' and 'non-differential interactions', i.e., interactions with or without differential changes in this work.

A differential interaction is an edge whose strength has significantly changed under different conditions (e.g., GK and WKY). On the other hand, a non-differential

interaction is an edge that has no significant change on edge strength but its linked two genes or corresponding coded proteins are both differentially expressed. Their specific schemes are shown in Figures 3E and 3F, respectively.

Co-expression using correlation coefficient is a popular method to describe the interaction or edge strength between genes, and the Spearman correlation coefficient was used here. Firstly, all genes were mapped to the compiled rat PPI network to filter some unnecessary interactions. As a result, 24,503 edges among 4,082 genes were selected. Then for each remaining edge, we separately computed its correlation coefficient under two contrast conditions. If it is strongly correlated (the Spearman correlation coefficient is greater than or equal to threshold 0.7) under one condition, while not in other conditions (the Spearman correlation coefficient is lower than 0.5), then this edge is defined as 'differential interactions'. With regard to non-differential interactions, they are detected based on differential expressions on its linked two genes. Specifically, if two endpoints (linked genes or coded proteins) of an edge are both differentially expressed, then this edge is noted as 'non-differential interactions'.

More importantly, from these detected 'non-differential interactions' and 'differential interactions', we are able to further detect 'disease interactions', in contrast to the traditional 'disease genes' (see Figure 1).

Measuring significant overlapping of two gene sets. We measured the significant overlapping of two gene sets based on the hyper-geometric probability. When one gene set is composed of k genes, and l genes are detected in the other gene set, the probability is obtained by

$$P(X \le l) = 1 - \sum_{i=0}^{l} \frac{\binom{M}{i}\binom{N-M}{k-i}}{\binom{N}{k}}$$

where M and N are the total number of genes in the two gene sets, respectively. In this study, we set the significance probability to 0.01.

Computing relative expression value. To characterize the differential expression levels between case and control, the relative gene expression was introduced and computed by normalizing the GK data on the basis of corresponding WKY value. Given one gene, its relative expression from the i_{th} GK sample at the t_{th} time point can be computed as follows:

$$relative_{i,t} = log2 \frac{GK_{i,t}}{\sum_{i=1}^{n} WKY_{i,t} \Big/ n}$$

In the above formula, n means the number of samples used in the t_{th} time point.

1. Frayling, T. M. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nat. Rev. Genet.* **8**, 657–662 (2007).
2. Prospective Diabetes Study Group U. K. Overview of 6 years' therapy of type II diabetes: a progressive disease. *Diabetes.* **44**, 1249–58 (1995).
3. Kostka, D. & Spang, R. Finding disease specific alterations in the co-expression of genes. *Bioinformatics.* **20**, i194–i199 (2004).
4. Ideker, T. & Krogan, N. J. Differential network biology. *Mol. Syst. Biol.* **8**, 565, doi:10.1038 (2012).
5. He, B., Zhang, H. & Shi, T. A comprehensive analysis of the dynamic biological networks in HCV induced hepatocarcinogenesis. *PLoS One.* **6**, e18516 (2011).
6. Ray, M. & Zhang, W. Analysis of Alzheimer's disease severity across brain regions by topological analysis of gene co-expression networks. *BMC. Syst. Biol.* **4**, 136 (2010).
7. Goni, J. *et al*. A computational analysis of protein-protein interaction networks in neurodegenerative diseases. *BMC. Syst. Biol.* **2**, 52 (2008).
8. Liu, M. *et al*. Network-based analysis of affected biological processes in type 2 diabetes models. *PLoS. Genet.* **3**, e96 (2007).
9. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, article 17 (2005).
10. Liu, Z. P., Wang, Y., Zhang, X. S & Chen, L. N. Network-based analysis of complex diseases. *IET. Syst. Biol.* **6**, 122–133 (2012).
11. Chen, L. N., Wang, R. S. & Zhang, X. S. Biomolecular Networks: Methods and Applications in Systems Biology. John Wiley & Sons, Hoboken, New Jersey (2009).
12. He, D. N., Liu, Z. P., Honda, M., Kaneko, S. & Chen, L. N. Coexpression network analysis in chronic hepatitis B and C hepatic lesion reveals distinct patterns of disease progression to hepatocellular carcinoma. *J. Mol. Cell. Biol.* **4**, 140–152 (2012).
13. Liu, Z. P., Wang, Y., Zhang, X. S., Xia, W. M. & Chen, L. N. Detecting and analyzing differentially activated pathways in brain regions of Alzheimer's disease patients. *Mol. Biosyst.* **7**, 1441–1452 (2011).
14. Fuller, T. F., Ghazalpour, A., Aten, J. E., Drake, T. A., Lusis, A. J. & Horvath, S. Weighted gene coexpression network analysis strategies applied to mouse weight. *Mamm. Genome.* **18**, 463–472 (2007).
15. Watson, M. Coexpress: differential co-expression in gene expression data. *BMC Bioinformatics.* **7**, 509–521 (2006).
16. de La Fuente, A. From 'differential expression' to 'differential networking' – identification of dysfunctional regulatory networks in diseases. *Cell.* **26**, 326–333 (2010).
17. Xue, B., Sukumaran, S., Nie, J., Jusko, W. J., Dubois, D. C. & Almon, R. R. Adipose tissue deficiency and chronic inflammation in diabetic Goto-Kakizaki rats. *PLoS One.* **6**, e17386 (2011).
18. Goldfine, I. D. *et al*. The role of membrane glycoprotein plasma cell antigen 1/ectonucleotide pyrophosphatase phosphodiesterase 1 in the pathogenesis of insulin resistance and related abnormalities. *Endocr. Rev.* **29**, 62–75 (2008).
19. Ramprasath, T., Senthi Murugan, P., Prabakaran, A. D., Gomathi, P., Rathinavel, A. & Sellvam, G. S. Potential risk modifications of GSTT1, GSTM1 and GSTP1 (glutathione-S-transferases) variants and their association to CAD in patients with type-2 diabetes. *Biochem. Biophys. Res. Commun.* **407**, 49–53 (2011).
20. Curtis, J. M. *et al*. Downregulation of adipose glutathione S-transferase A4 leads to increased protein carbonylation, oxidative stress, and mitochondrial dysfunction. *Diabetes.* **59**, 1132–1142 (2010).
21. Tong, Y. C. & Cheng, J. T. Changes in bladder nerve-growth factor and p75 genetic expression in streptozotocin-induced diabetic rats. *BJU. Int.* **96**, 1392–1396 (2005).
22. Kim, S. K., Abdelmegeed, M. A. & Novak, R. F. Identification of the Insulin Signaling Cascade in the Regulation of Alpha-Class Glutathione *S*-Transferase Expression in Primary Cultured Rat Hepatocytes. *J. Pharmacol. Exp. Ther.* **316**, 1255–1261 (2006).
23. Herrera, B. M. *et al*. Global microRNA expression profiles in insulin target tissues in a spontaneous rat model of type 2 diabetes. *Diabetelogia.* **53**, 1099–1109 (2010).
24. Thevenod, F. Pathophysiology of diabetes mellitus type 2: roles of obserity, insulin resistance and β-cell dysfunction. In: Masur, K., Thévenod, F., Zänker, K. S. (eds) Diabetes and Cancer (Epidemiological Evidence and Molecular Links), Karger, Basel. **19**, 1–18 (2008).
25. Kido, Y. *et al*. Tissue-specific insulin resistance in mice with mutations in the insulin receptor, IRS-1, and IRS-2. *J. Clin. Invest.* **105**, 199–205 (2000).
26. Fritsche, L. *et al*. Insulin-induced serine phosphorylation of IRS-2 via ERK1/2 and mTOR: studies on the function of Ser675 and Ser907. *Am. J. Physiol. Endocrinol. Metab.* **300**, 824–836 (2011).
27. Quon, M. J. *et al*. Roles of 1-phosphatidylinositol 3-kinase and ras in regulating translocation of GLUT4 in transfected rat adipose cells. *Mol. Cell. Biol.* **15**, 5403–5411 (1995).
28. Abel, E. D. *et al*. Adipose-selective targeting of the GLUT4 gene impairs insulin action in muscle and liver. *Nature.* **409**, 729–733 (2001).
29. Rondinone, C. M., Wang, L. M., Lonnroth, P., Wesslau, C., Pierce, J. H. & Smith, U. Insulin receptor substrate (IRS) 1 is reduced and IRS-2 is the main docking protein for phosphatidylinositol 3-Kinase in adipocytes from subjects with non-insulin-dependent diabetes mellitus. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 4171–4175 (1997).
30. Sciacchitano, S. & Taylor, S. I. Cloning, tissue expression, and chromosomal localization of the mouse IRS-3 gene. *Endocrinology.* **138**, 4931–4940 (1997).
31. Zhou, L. *et al*. Action of insulin receptor substrate-3 (IRS-3) and IRS-4 to stimulated translocation of GLUT4 in rat adipose cells. *Mol. Endocrinol.* **13**, 505–514 (1999).
32. Sesti, G., Federici, M., Hribal, M. L., Lauro, D., Sbraccia, P. & Lauro, R. Defects of the insulin receptor substrate (IRS) system in human metabolic disorders. *FASEB. J.* **15**, 2099–2111 (2001).
33. Sandu, O. A., Ragolia, L. & Begum, N. Diabetes in the Goto-Kakizaki rat is accompanied by impaired insulin-mediated myosin-bound phosphatase activation and vascular smooth muscle cell relaxation. *Diabetes.* **49**, 2178–2189 (2000).
34. Almon, R. R., Dubois, D. C., Lai, W., Xue, B., Nie, J. & Jusko, W. J. Gene expression analysis of hepatic roles in cause and development of diabetes in Goto-Kakizaki rats. *J. Endocrinol.* **200**, 331–346 (2009).
35. Zhong, Q. *et al*. Edgetic perturbation models of human inherited disorders. *Mol Syst. Biol.* **5**, 321 (2009).
36. Liu, X. P., Liu, Z. P., Zhao, X. M. & Chen, L. N. Identifying disease genes and module biomarkers with differential interactions. *J. Am. Med. Inform. Assoc.* **19**, 241–248 (2011).
37. Chen, L. N., Liu, R., Liu, Z. P., Li, M. Y. & Aihara, K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci. Rep.* **2**, 342 (2012).
38. Chuang, H., Lee, E., Liu, Y., Lee, D. & Ideker, T. Network-based classification of breast cancer metastasis. *Mol. Syst. Biol.* **3**, 140 (2007).
39. Nie, J., Xue, B., Sukumaran, S., Jusko, W. J., Dubois, D. C. & Almon, R. R. Differential muscle gene expression as a function of disease progression in Goto-Kakizaki diabetic rats. *Mol. Cell. Endocrinol.* **338**, 10–17 (2011).
40. Zhao, J., Yang, T., Huang, Y. & Holme, P. Ranking candidate disease genes from gene expression and protein interaction: a katz-centrality based approach. *PLoS One.* **6**, e24306, doi:10.1371/journal.pone.0024306 (2011).
41. Corkey, B. E. Banting lecture 2011: hyperinsulinemia: cause or consequence? *Diabetes.* **61**, 4–13 (2012).
42. Kalupahana, N. S. & Moustaid-Moussa, N. The renin-angiotensin system: a link between obesity, inflammation and insulin resistance. *Obes. Rev.* **13**, 136–49 (2012).

43. Schäfer, S. A., Machicao, F., Fritsche, A., Häring, H. U. & Kantartzis, K. New type 2 diabetes risk genes provide new insights in insulin secretion mechanisms. *Diabetes. Res. Clin. Pract.* **93**, S9–24 (2011).

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) with Nos. 61134013, 61072149, 91029301, 61171007, 31100949 and 11131009. This work was also supported by the Chief Scientist Program of Shanghai Institutes for Biological Sciences of CAS with the grant number 2009CSP002, by the National Center for Mathematics and Interdisciplinary Sciences of CAS, by the Knowledge Innovation Program of CAS with grant No. KSCX2-EW-R-01, and by the FIRST program from JSPS initiated by CSTP.

Author contributions

Y.W. and L.C. designed research; S.S., Z.L., T.Z. and Y.W. performed the research and analyzed data; S.S., Y.W. and L.C. wrote the paper; and all authors reviewed and revised the paper.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Sun, S., Liu, Z., Zeng, T., Wang, Y. & Chen, L. Spatio-temporal analysis of type 2 diabetes mellitus based on differential expression networks. *Sci. Rep.* **3**, 2268; DOI:10.1038/srep02268 (2013).