METABOLOMICS DIFFERENTIAL CORRELATION NETWORK ANALYSIS OF OSTEOARTHRITIS

Ting Hu*

Department of Computer Science
Memorial University, St. John's, NL, Canada
**E-mail: ting.hu@mun.ca*

Weidong Zhang

Discipline of Genetics, Faculty of Medicine
Memorial University, St. John's, NL, Canada

Zhaozhi Fan

Department of Mathematics and Statistics
Memorial University, St. John's, NL, Canada

Guang Sun

Discipline of Medicine, Faculty of Medicine
Memorial University, St. John's, NL, Canada

Sergei Likhodi

Department of Laboratory Medicine, Faculty of Medicine
Memorial University, St. John's, NL, Canada

Edward Randell

Department of Laboratory Medicine, Faculty of Medicine
Memorial University, St. John's, NL, Canada

Guangju Zhai

Discipline of Genetics, Faculty of Medicine
Memorial University, St. John's, NL, Canada

Osteoarthritis (OA) significantly compromises the life quality of affected individuals and imposes a substantial economic burden on our society. Unfortunately the pathogenesis of the disease is till poorly understood and no effective medications have been developed. OA is a complex disease that involves both genetic and environmental influences. To elucidate the complex interlinked structure of metabolic processes associated with OA, we developed a differential correlation network approach to detecting the interconnection of metabolite pairs whose relationships are significantly altered due to the diseased process. Through topological analysis of such a differential network, we identified key metabolites that played an important role in governing the connectivity and information flow of the network. Identification of these key metabolites suggests the association of their underlying cellular processes with OA and may help elucidate the pathogenesis of the disease and the development of novel targeted therapies.

Keywords: Differential correlation; Osteoarthritis; Metabolomics; Urea cycle abnormal; Obesity; Cardiovascular diseases; Differential networks; Dynamical networks; Interaction mapping.

1. Introduction

Osteoarthritis (OA) is the most common form of arthritis. It causes a substantial morbidity and disability in the elderly populations, and imposes a great economic burden on our society.[1,2] Despite high prevalence and societal impact, there is no medication that can cure it, or reverse or halt the disease progression, partly because that its pathogenesis is still unclear and there is no reliable method that can be used for early OA diagnosis.

Recent developments in the field of metabolomics provide an array of new tools for the study of OA. A large number of small-molecule metabolites from body fluids or tissues can be quantitatively detected simultaneously, which promises an immense potential for early diagnosis, therapy monitoring and understanding the pathogenesis of complex diseases.[3] Metabolites are intermediate and end products of various cellular processes and their levels of concentration serve as a good indicator of a sequence of biological systems in response to genetic and environmental influences.

In the reported studies on metabolomics analysis of OA case-control population data, the mostly adopted methodology is to test and identify metabolites that are significantly associated with the disease class using principal component analysis (PCA),[4,5] partial least square discriminant analysis (PLS),[6,7] or other individual testing techniques, and then to deduce their likely biological interrelationship with OA. Testing correlations of the concentrations of metabolites has not seen wide adoption likely due to the limited availability of methodologies. However, these correlations likely exist because metabolites are intermediate or end products of interconnected cellular processes. Analyzing their correlations provides an avenue capturing the relationships of their represented cellular processes and biological reactions associated with OA, and thus holds a great potential in OA metabolomics research.

Meanwhile, many biological systems are increasingly viewed and analyzed as highly complex networks of interlinked molecular or cellular entities or metabolites,[8] and network science has been applied to capture the interactome maps of gene-gene or protein-protein interactions[9–13] as well as transcriptional and metabolic data.[14–16]

The interaction maps of proteins, genes, metabolites or diseases can reveal the overall physical and functional landscape of a biological system, and these networks have been mostly generated under a particular static condition. More recently, differential network analysis has been promoted as a powerful framework for analyzing biological interaction maps when biological systems are considered undergoing differential changes that are dependent on the environment, tissue type, disease state, development or speciation.[17,18]

Recent interaction mapping studies have demonstrated the power of differential correlation analysis for elucidating the re-wiring of the interaction architecture of fundamental biological responses in adaptation to changing conditions.[19–25] Analyzing the rewiring of biological networks across disease conditions provides a unique insight into the dynamic response of a biological system. Instead of looking at the absolute properties of a system, differential network analysis emphasizes on the characteristics that are the most affected by genetic or environmental influences.

In this study, we proposed a differential network approach to analyzing the metabolomics population-based data of OA. We used differential analysis to quantify the variation of pair-

wise correlation of metabolites across case and control populations, and used networks to characterize the global interconnecting structure of such differentially correlated metabolites. Our methodology is distinct from most existing metabolomics analyses of OA in that we investigated the correlations of metabolite concentrations, and more importantly the variations of such correlations by comparing different disease status, to help elucidate the underlying biological processes specifically associated with the pathogenesis of OA. Using topological analysis of such a differential correlation network, we identified key metabolites and subsequently their represented cellular processes that may play an important role in the clinical development of OA. Our findings could be very helpful in designing novel and more targeted therapies for OA.

2. Methods

2.1. *Osteoarthritis metabolomics data*

In the current study, we used a two-stage case-control design with a discovery phase and a validation phase. For both phases, knee OA patients were selected from the Newfoundland Osteoarthritis Study (NFOAS) initiated in 2011.[26] The NFOAS aimed at identifying novel genetic, epigenetic, and biochemical markers for OA. The NFOAS recruited OA patients who underwent a total knee replacement surgery due to primary OA between November 2011 and December 2013 at the St. Clare's Mercy Hospital and Health Science Centre General Hospital in St. John's, the capital city of Newfoundland and Labrador (NL), Canada. Healthy controls for both phases were selected from the CODING study (The Complex Diseases in the Newfoundland population: Environment and Genetics), where participants were adult volunteers.[27]

Both cases and controls were from the same source population of Newfoundland and Labrador. Knee OA diagnosis was made based on the American College of Rheumatology clinical criteria for classification of idiopathic OA of the knee[28] and the judgment of the attending orthopedic surgeons. Controls were individuals without self-reported family doctor diagnosed knee OA based on their medical information collected by a self-administered questionnaire. We collected 64 OA cases and 45 healthy controls in the discovery phase and 72 cases and 76 controls in the replication phase.

Blood samples were collected after at least 8 hour fasting and plasma was separated from blood using the standard protocol. Metabolic profiling was performed on plasma using the Waters XEVO TQ MS system (Waters Limited, Mississauga, Ontario, Canada) coupled with Biocrates AbsoluteIDQ p180 kit, which measures 186 metabolites including 90 glycerophospholipids, 40 acylcarnitines (1 free carnitine), 21 amino acids, 19 biogenic amines, 15 sphingolipids and 1 hexose (above 90 percent is glucose). The details of the 186 metabolites and the metabolic profiling method were described in the previous publication.[29] Over 90% of the metabolites (167/186) were successfully determined in each sample.

Age and BMI are known factors correlated with OA. Therefore, the residual of a linear regression using attributes age and BMI was applied to remove any partial correlations as a result of those two factors, and to adjust the data for our metabolomics differential correlation analysis of OA.

2.2. *Differential analysis of metabolite correlations*

Metabolite concentrations in plasma may be correlated as a result of their represented biological processes, and the correlation may change in different phenotypic or disease conditions. Such a dynamic correlation was quantified by a differential correlation statistic in our study.

The correlation of a pair of metabolites was calculated using Pearson's correlation coefficient r in the two phenotypically distinguished samples, i.e. cases and controls. The correlation coefficients r_{case} and r_{control} were then used to compute the change of the correlation between two metabolites across two different disease classes. Specifically, for metabolites i and j, their differential correlation $r_{\text{diff}}(i,j)$ is calculated as the normalized difference of Fisher's z-transformations of $r_{\text{case}}(i,j)$ and $r_{\text{control}}(i,j)$,

$$r_{\text{diff}}(i,j) = \sqrt{\frac{n_{\text{case}}-3}{2}} \times z_{\text{case}}(i,j) - \sqrt{\frac{n_{\text{control}}-3}{2}} \times z_{\text{control}}(i,j), \tag{1}$$

where z is the Fisher's z-transformation of correlation coefficient r,

$$z_{\text{case}}(i,j) = \frac{1}{2}\ln\big[\frac{1+r_{\text{case}}(i,j)}{1-r_{\text{case}}(i,j)}\big],\ z_{\text{control}}(i,j) = \frac{1}{2}\ln\big[\frac{1+r_{\text{control}}(i,j)}{1-r_{\text{control}}(i,j)}\big]. \tag{2}$$

We used n_{case} and n_{control} to denote the total numbers of samples in cases and controls. This differential correlation statistic captures the change of the normalized correlation across two distinguishing conditions, and we used it to test if two metabolites are differentially correlated by comparing diseased and healthy populations. Note that r_{diff} describes the change of correlations by subtracting the correlation in controls from that in cases, and can take either positive or negative values.

The significance levels of differential correlations were assessed using a 1000-fold permutation test. For each permutation, we randomly shuffled the disease status of all samples combining both cases and controls to remove the association among metabolite correlations and the disease outcome. By repeating this process 1000 times, we were able to generate a null distribution under the assumption that the pairwise correlations of metabolites were not statistically distinguishing in cases and in controls. Then for each pair of metabolites, the significance (p-value) of their differential correlation was estimated as the proportion of permuted differential correlations that were greater than the observed value calculated using the original real data.

2.3. *Differential correlation network*

Network is a powerful tool to characterize the properties of entities and their complex relationships. In this study, we used networks to represent the global structure of differentially correlated metabolites by comparing OA cases and healthy controls.

Pairs of metabolites that had significant differential correlations were included to build the network. In such a differential correlation network, each node stood for a metabolite, and edges linking two metabolites represented the significant differential correlations between them. The differential correlation of a metabolite pair could be either positive or negative, meaning that their correlation in cases are significantly stronger than their correlation in controls or vice versa.

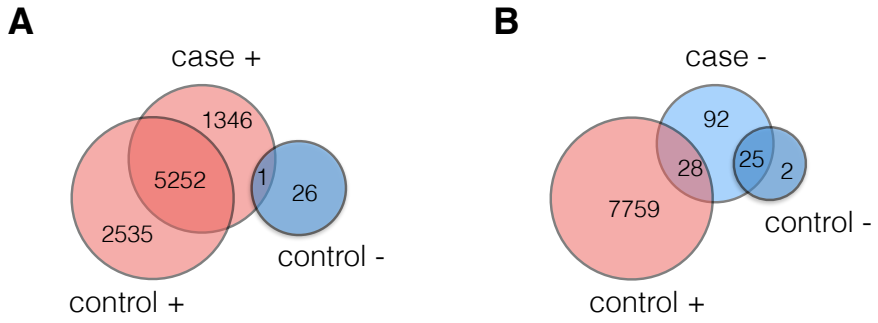

Fig. 1. Comparison of pairwise metabolite correlations (red for positive; blue for negative) in case and control populations. Only significant correlations (Pearson's correlation coefficient p-value cutoff 0.05 with Bonferroni multiple-testing correction) were included in this comparison. (**A**) For the total of 6599 ($= 1346 + 5252 + 1$) pairs of positively correlated metabolites in cases, the majority of them were also found positively correlated in controls. (**B**) For the total of 145 ($= 92 + 28 + 25$) pairs of negatively correlated metabolites in cases, a third of them were found positively correlated in controls and another third of them were found negatively correlated in controls.

3. Results

3.1. *Metabolite correlations in case and control populations*

The pairwise Pearson's correlations of 167 metabolites were calculated in both case and control samples in the discovery dataset. Of all 13,861 pairs, the majority of them were positively correlated in both cases and controls. We used a p-value threshold 0.05 and Bonferroni multiple-testing correction to define the statistical significance of pairwise correlations.

About 80% of the positively correlated pairs in cases were found also positively correlated in controls (Fig. 1**A**), and a similar link was observed for negatively correlated pairs as well (Fig. 1**B**). This large overlapping of metabolite correlations from the two phenotypic conditions suggests that the majority of the observed correlations were a result of "housekeeping" biological reactions and were not related to the disease of OA.

3.2. *Differentially correlated metabolites*

We calculated the differential correlations of all pairs of metabolites by comparing their correlations in cases and controls as described in the section of Methods. By subtracting correlations in controls from correlations in cases, metabolite pairs that were differentially correlated across these two conditions were magnified, while the persistent correlations in both conditions were removed. This differential correlation method allowed us to focus on the dynamic correlations that were specifically associated with the disease.

In the discovery dataset, 232 pairs of metabolites had significant positive differential correlations and 1060 pairs had significant negative differential correlations (permutation testing $p < 0.05$). The strongest and most significant pair of metabolites that has a positive differential correlation is Ala and Sarcosine ($r_{\text{diff}} = 9.33$, $p < 0.001$), and that has a negative differential correlation is lysoPCaC24:0 and PCaaC40:2 ($r_{\text{diff}} = -5.40$, $p < 0.001$). Fig. 2**A** shows a scatter plotting of all the pairs of metabolites with their correlations in the case population (x-axis) and the control population (y-axis). In addition, positive and negative differential correla-

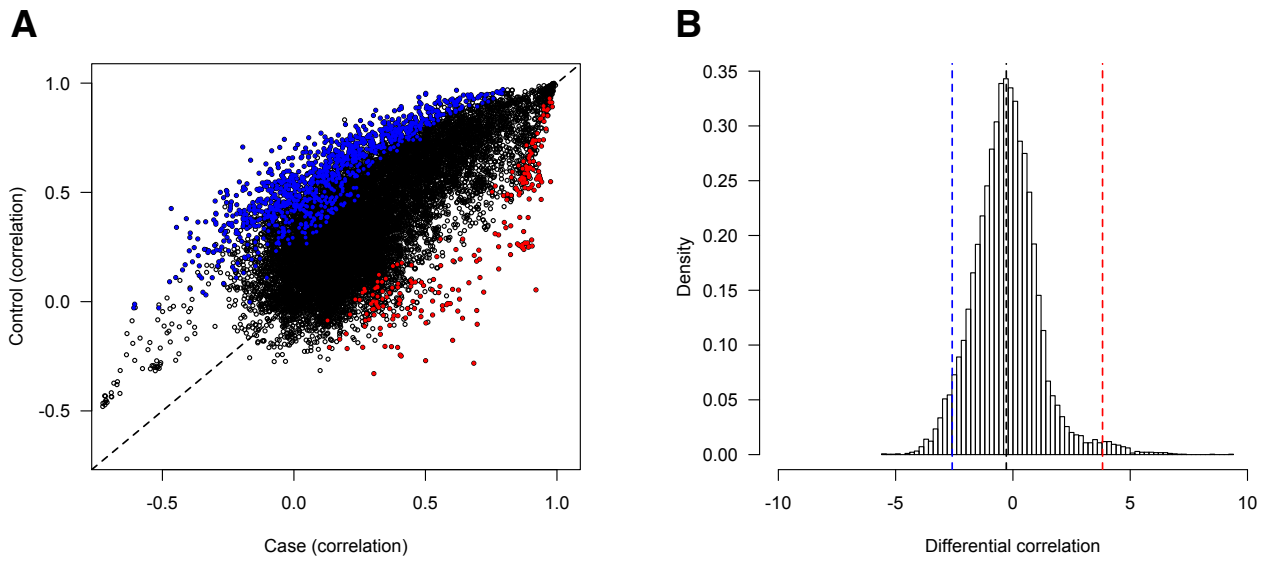

Fig. 2. (**A**) Scatter of metabolite pair correlations in cases (x axis) and controls (y axis) and identification of significant (p-value cutoff 0.05 using a 1000-fold permutation test) pairs with positive differential correlations (red) and with negative differential correlations (blue). (**B**) Distribution of all pairwise differential correlations, with a mean value of -0.283 (black dashed line). The means of significant differential correlations are also shown using dashed lines. The average significant positive correlations was 3.820 (red) and the average significant negative correlations was -2.589 (blue).

tions were shown as colored points. They represented the metabolite pairs whose correlations significantly changed across the two phenotypic conditions.

The distribution of the differential correlations of all metabolite pairs is shown in Fig. 2**B**. It follows a normal distribution approximately with a mean of -0.283. The shift of this distribution towards the negative values explained the observation that there were more significant negative differential correlations (1060 pairs) than positive ones (232 pairs). However, positive differential correlation distribution has a longer tail towards larger values, and the mean of significant positive differential correlations, i.e. 3.820, was greater than the absolute of the mean of negative ones, i.e. -2.589.

3.3. *Differential correlation network of OA*

We applied differential correlation analysis to both the discovery and replication datasets. We used the set of metabolite pairs that were significantly differentially correlated (permutation testing significance cutoff $p < 0.05$) in both datasets to build the differential correlation network of OA.

A total of 100 pairwise differential correlations were statistically significant in both datasets, including 71 metabolites. The network was comprised of four connected components and the largest component included 63 metabolites and 95 edges (Fig. 3). The remaining three components had only two or four nodes and were not included in the network visualization.

As seen in the figure, the majority of metabolite pairs were negatively differentially correlated, denoted by blue edges in the graph. Positive differential correlations, however, were

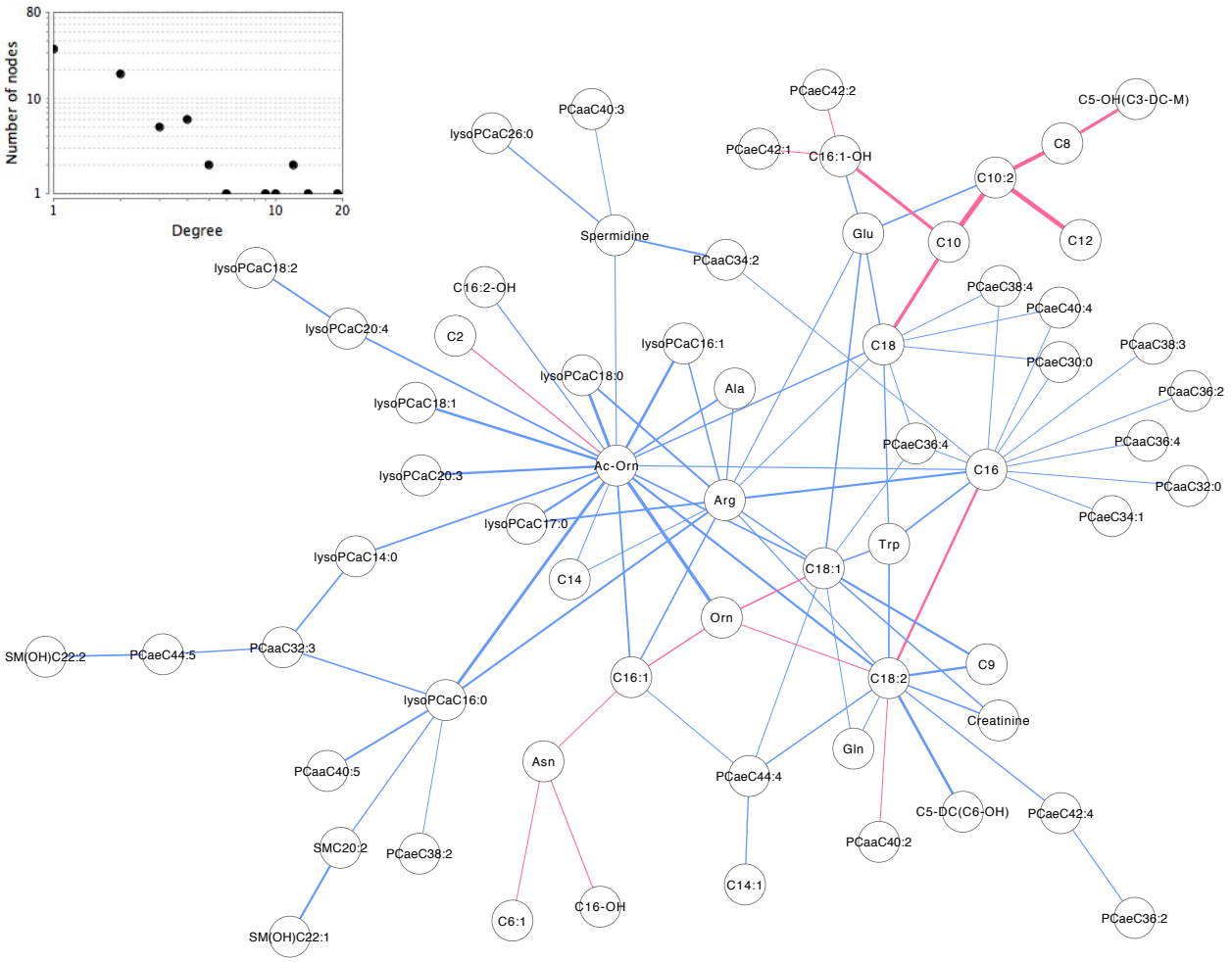

Fig. 3. The differential correlation network by comparing the discovery and replication data. Only pairs of metabolites that have significant differential correlations in both datasets are shown. There is one major connected component of the network, which has 63 nodes and 95 edges. The network is visualized using the force-directed layout with a closer node layout distance representing a stronger pairwise correlation. Edge width is proportional to the corresponding correlation strength and edge color codes for positive (red) and negative (blue) differential correlations. This network visualization was generated using Cytoscape.[30]

less observed and clustered together in sub-structures of the network. The node degree of this network had a mean of 3.02 and a heavy-tail distribution (inset of Fig. 3), showing that the majority of nodes have a very low degree but a few of them were considerably more connected than the others. This property suggests the robustness of connectivity and information flow in the network.

3.4. *Identification of key metabolites in the osteoarthritis differential correlation network*

In network science, the importance of an individual node in a network is captured by measuring its *centrality*. Besides the most commonly used centrality measure, node degree, there are more sophisticated metrics on node importance that characterize not only the number of connections

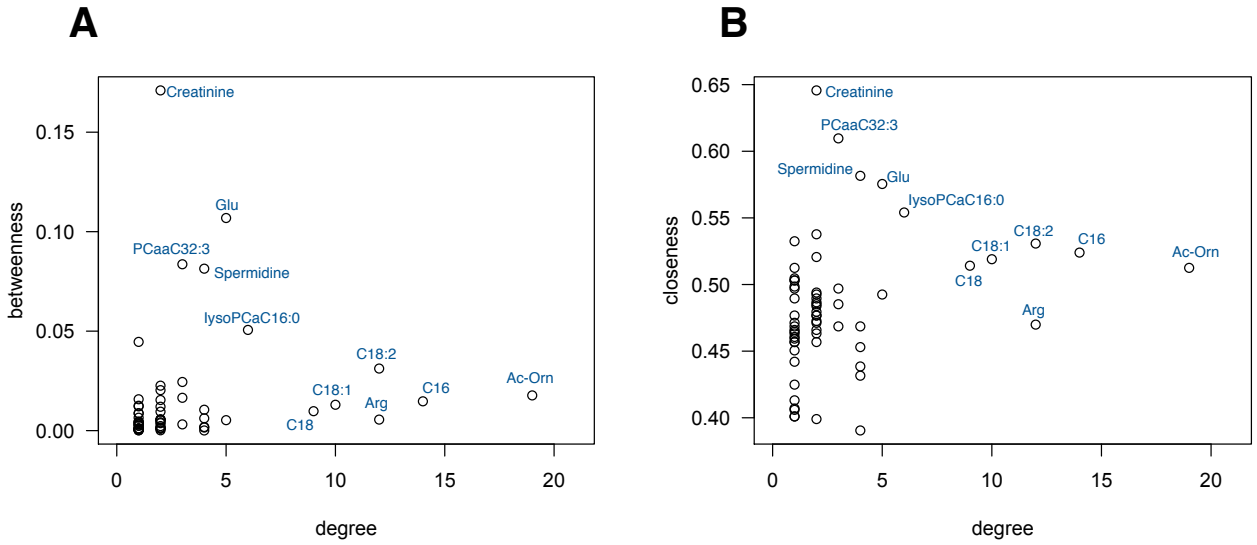

Fig. 4. Node importance characterized by (**A**) betweenness centrality and (**B**) closeness centrality in relation to node degree. Key nodes, either with high degrees, or high betweenness/closeness, or both, are identified and labeled with their represented metabolite names.

a node has, but also on how important those connections are in the global structure of an entire network. *Betweenness* centrality quantifies the number of times a node v is part of the shortest path between any pair of nodes,[31] represented as $\sum_{s\neq v\neq t\in V}\frac{\sigma_{st}(v)}{\sigma_{st}}$, where σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}(v)$ is the number of those paths that pass through node v. Betweenness captures how important a given node acts on the connectivity of all other pairs of nodes. *Closeness* centrality is defined as $\frac{1}{\sum_{s\neq v} d_{vs}}$, where d_{vs} is the distance between nodes v and s.[32,33] This metric describes how easily a given node can reach all other nodes in a network. In the context of differential correlation networks, those centrality measures were used to identify key metabolites that play an essential role in the global interconnected structure.

Nodes with high degrees are usually referred to as "hubs" since they have more connections than the rest of the nodes in the network, and nodes with high betweenness or closeness are often referred to as "bottlenecks" since they are crucial in controlling the information flow in the network. Fig. 4 shows metabolites that are hubs, or bottlenecks, or both. The betweenness and closeness centralities are shown in relation to node degrees in the figure. The same set of 11 metabolites were identified as key nodes in both centrality measures (Fig. 4**A** and **B**).

4. Discussion

Identification of metabolic markers associated with OA holds a great potential to better understand the cellular processes in response to genetic and environmental influences that lead to the clinical outcome of the disease. The identified metabolites and their represented cellular processes will in turn help us to develop targeted therapies for OA. In this study, we developed a differential network approach to characterizing the variations of metabolite correlations in relation to different phenotypic conditions.

In our methodology, we used networks to represent the global inter-connected structure of metabolites that showed significant correlation variations in case and control populations. By exploring the topological properties of such a differential correlation network, we identified a set of key metabolites for modulating connectivity and information flow in the network, and thus hypothesized the association of their represented cellular processes with the disease.

When metabolite correlations were analyzed separately in cases and controls, we saw a large overlap of correlated metabolite pairs (Fig. 1), an observation indicating that most of the metabolite associations are not specifically related to OA. The differential analysis took a unique route by subtracting correlation coefficient of a metabolite pair in controls from that in cases, such that all the persistent pairwise correlations across the two phenotypic classes were removed and the pairs with significant variations were magnified. These differentially correlated metabolites are expected to provide useful insights into the underlying biological processes of the clinical development of OA. We observed considerably more significant negative differential correlations than positive ones (Fig. 2), which indicates that important biological processes might be compromised in OA patients.

By comparing the independent discovery and replication datasets, we built a differential correlation network of metabolites associated to OA (Fig. 3) The network included 63 metabolites and 95 pairwise differential correlations. The majority of the differential correlations were negative while the positive ones were clustered together around certain metabolites. The metabolites that have positive differential correlations are mainly coming from the same class of acylcarnitines, e.g. C18, C10, C10:2, C8, C5-OH(C3-DC-M), C12 and C16:1-OH; C18:2 and C16; C6:1, C16-OH and C16:1. From the view point of physiology function, the relationship between these metabolites is more likely a parallel relation rather than a causality.

The node degrees of this differential network had a heavy tail distribution (Fig. 3 inset), which suggests a robust property of connectivity and information flow subject to random perturbations. That is, random removal of nodes will have a very limited impact on the global connectivity of the network, a property that has been found in many biological systems including metabolic networks,[14] protein-protein interaction networks,[34] gene-regulatory networks[8] and gene-gene interaction networks.[35] In the context of OA metabolite differential correlation networks, this robustness property indicates the complexity of the molecular and cellular processes underlying the pathogenesis of OA.

Topological analysis on the node importance using centrality measures revealed a set of key metabolites that play an essential role modulating the connectivity and information flow in the network (Fig. 4). They were identified as "hubs", i.e. nodes that connect to many other nodes, and "bottlenecks", i.e. nodes that are located on major information flow paths in the network. Identification of these key metabolites may provide important insights into the pathogenesis of OA. Based on the node centrality measures, the metabolites in the network can be roughly classified into three categories. The hub-and-bottleneck metabolites Ac-Orn and Arg with their close neighbors Ala and Orn comprise the core of the network. On the network peripheral, metabolites are mostly glycerophospholipids (PC and LysoPC). Between the core and peripheral of the network is where acylcarnitines mixed with glycerophospholipids are located.

Ac-Orn, Arg, Ala and Orn have a close relationship with urea cycle in the body. Previous studies have proposed that urea cycle disorders may be related to the OA pathogenesis.[36,37] Glycerophospholipids form the essential lipid bilayer of all biological membranes and are closely involved in signal transduction, regulation of membrane trafficking and many other membrane-related phenomena.[38,39] It has been suggested that alterations in phospholipid composition and concentrations are associated with the development of OA.[40]

Acylcarnitines are related to energy metabolism. Carnitine and its acyl esters acylcarnitines are essential compounds for the metabolism of fatty acids. Carnitine can assist in the transport and metabolism of fatty acyl-CoA from the cytosol to the mitochondrial matrix, where the enzymes of oxidation are located and fatty acids are oxidized as a major source of energy. Acylcarnitine abnormal have been detected in obesity, type-2 diabetes, and cardiovascular diseases.[41,42]

The clustering of metabolites in the differential correlation network based on their centralities and the observation of urea cycle related metabolites locating on the core cluster of the network suggest that urea cycle abnormality may be a driving cause for metabolic disorders and may have a significant influence on OA development.

Acknowledgements

This work is partially supported by Newfoundland and Labrador Research & Development Corporation (RDC) Ignite grant to TH. ZF is supported by the Natural Science and Engineering Research Council (NSERC) of Canada. GZ acknowledges grants from Canadian Institute of Health Research (CIHR), Newfoundland and Labrador Research & Development Corporation (RDC) and Memorial University. We thank all the study participants who made this study possible and all the Operation Room staff in Eastern Health General Hospital and St. Clare's Hospital who helped for collecting samples.

References

1. WHO Scientific Group, *WHO Technical Report Series* **919**, p. 218 (2003).
2. J. Y. Reginster, *Rheumatology* **41**, 3 (2004).
3. G. N. Gowda, S. Zhang, H. Gu, V. Asiago, N. Shanaiah and D. Raftery, *Expert Review of Molecular Diagnostics* **8**, 716 (2008).
4. K. J. Deluzio and J. L. Astephen, *Gain and Posture* **25**, 86 (2007).
5. B. J. de Lange-Brokaar, A. Ioan-Facsinay, E. Yusuf, A. W. Visser, H. M. Kroon, G. J. van Osch, A. M. Zuurmond, V. Stojanovic-Susulic, J. L. Bloem, R. G. Nelissen, T. W. Huizinga and M. Kloppenburg, *Arthritis and Rheumatology* **67**, 733 (2015).
6. L. M. Gierman, S. Wopereis, B. van El, E. R. Verheij, B. J. Werff-van der Vat, Y. M. Bastiaansen-Jenniskens, G. J. van Osch, M. Kloppenburg, V. Stojanovic-Susulic, T. W. Huizinga and A. M. Zuurmond, *Arthritis and Rheumatology* **65**, 2606 (2013).
7. C. Wu, R. Lei, M. Tiainen, S. Wu, Q. Zhang, F. Pei and X. Guo, *Experimental Cell Research* **326**, 240 (2014).
8. A.-L. Barabasi and Z. N. Oltvai, *Nature Review Genetics* **5**, 101 (2004).
9. P. F. Jonsson and P. A. Bates, *Bioinformatics* **22**, 2291 (2006).
10. O. Vanunu, O. Magger, E. Ruppin, T. Shlomi and R. Sharan, *PLoS Computational Biology* **6**, p. e1000641 (2010).

11. S. Navlakha and C. Kingsford, *Bioinformatics* **26**, 1057 (2010).
12. T. Hu, Y. Chen, J. W. Kiralis and J. H. Moore, *Genetic Epidemiology* **37**, 283 (2013).
13. T. Hu, Q. Pan, A. S. Andrew, J. M. Langer, M. D. Cole, C. R. Tomlinson, M. R. Karagas and J. H. Moore, *BioData Mining* **7**, p. 5 (2014).
14. H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai and A. L. Barabasi, *Nature* **407**, 651 (2000).
15. E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai and A. L. Barabasi, *Science* **297**, 1551 (2002).
16. P. Braun, E. Rietman and M. Vidal, *Proceedings of the National Academy of Sciences* **105**, 9849 (2008).
17. A. de la Fuente, *Trends in Genetics* **26**, 326 (2010).
18. T. Ideker and N. J. Krogan, *Molecular Systems Biology* **8**, p. 565 (2012).
19. Q. Zhong, N. Simonis, Q.-R. Li, B. Charloteaux, F. Heuze, N. Klitgord, S. Tam, H. Yu, K. Venkatesan, D. Mou, V. Swearingen, M. A. Yildirim, H. Yan, A. Dricot, D. Szeto, C. Lin, T. Hao, C. Fan, S. Milstein, D. Dupuy, R. Brasseur, D. E. Hill, M. E. Cusick and M. Vidal, *Molecular Systems Biology* **5**, p. 321 (2009).
20. I. W. Taylor, R. Linding, D. Warde-Farley, Y. Liu, C. Pesquita, D. Faria, S. Bull, T. Pawson, Q. Morris and J. L. Wrana, *Nature Biotechnology* **27**, 199 (09).
21. S. Bandyopadhyay, M. Mehta, D. Kuo, M.-K. Sung, R. Chuang, E. J. Jaehnig, B. Bodenmiller, K. Licon, W. Copeland, M. Shales, D. Fiedler, J. Dutkowski, A. Guenole, H. van Attikum, K. M. Shokat, R. D. Kolodner, W.-K. Huh, R. Aebersold, M.-C. Keogh, N. J. Krogan and T. Ideker, *Science* **330**, 1385 (2010).
22. J.-H. Chu, R. Lazarus, V. J. Carey and B. A. Raby, *BMC System Biology* **5**, p. 89 (2011).
23. B. Valcarcel, P. Wurtz, N.-K. S. al Basatena, T. Tukiainen, A. J. Kangas, P. Soininen, M.-R. Jarvelin, M. Ala-Korpela, T. M. Ebbels and M. de Iori, *PLoS One* **6**, p. e24702 (2011).
24. A. Fukushima, T. Nishizawa, M. Hayakumo, S. Hikosaka, K. Saito, E. Goto and M. Kusano, *Genome Analysis* **158**, 1487 (2012).
25. N. M. Penrod and J. H. Moore, *BMC Medical Genomics* **6**, p. S2 (2013).
26. G. Zhai, E. Aref-Eshghi, P. Rahman, H. Zhang, G. Martin, A. Furey, R. C. Green and G. Sun, *Journal of Orthopedics and Rheumatology* **1**, p. 5 (2014).
27. B. Fontaine-Bisson, J. Thorburn, A. Gregory, H. Zhang and G. Sun, *The American Journal of Clinical Nutrition* **99**, 384 (2014).
28. R. Altman, G. Alarcon, D. Appelrouth, D. Bloch, D. Borenstein, K. Brandt, C. Brown, T. D. Cooke and et al., *Arthritis and Rheumatology* **34**, 505 (1991).
29. W. Zhang, S. Likhodii, E. Aref-Eshghi, Y. Zhang, P. E. Harper, E. Randell, R. Green, G. Martin, A. Furey, G. Sun, P. Rahman and G. Zhai, *The Journal of Rheumatology* **42**, 859 (2015).
30. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, *Genome Research* **13**, 2498 (2003).
31. L. C. Freeman, *Sociometry* **40**, 35 (1977).
32. A. Bavelas, *Journal of the Acoustical Society of America* **22**, 725 (1950).
33. G. Sabidussi, *Psychometrika* **31**, 581 (1966).
34. H. Jeong, S. P. Mason, A. L. Barabasi and Z. N. Oltvai, *Nature* **411**, 41 (2001).
35. T. Hu, N. A. Sinnott-Armstrong, J. W. Kiralis, A. S. Andrew, M. R. Karagas and J. H. Moore, *BMC Bioinformatics* **12**, p. 364 (2011).
36. S. B. J. Adams, L. A. Setton, E. Kensicki, M. P. Bolognesi, A. P. Toth and D. L. Nettles, *Osteoarthritis Cartilage* **20**, 64 (2012).
37. S. Kim, J. Hwang, J. Xuan, Y. H. Jung, H.-S. Cha and K. H. Kim, *PLoS One* **9**, p. e97501 (2014).
38. M. J. Berridge and R. F. Irvine, *Nature* **341**, 197 (1989).
39. A. A. Farooqui, L. A. Horrocks and T. Farooqui, *Chemistry and Physics of Lipids* **106**, 1 (2000).

40. B. A. Hills, *Internal Medicine Journal* **32**, 242 (2002).
41. V. A. Zammita, R. R. Ramsayb, M. Bonominic and A. Arduini, *Advanced Drug Delivery Reviews* **61**, 1353 (2009).
42. M. A. Ramos-Roman, L. Sweetman, M. J. Valdez and E. J. Parks, *Metabolism Clinical and Experimental* **61**, 202 (2012).