



Xiao, Xiaolin and Dawson, Neil and MacIntyre, Lynsey and Morris, Brian J. and Pratt, Judith A. and Watson, David G. and Higham, Desmond J (2011) Exploring metabolic pathway disruption in the subchronic phencyclidine model of schizophrenia with the generalized singular value decomposition. BMC Systems Biology, 5. ISSN 1752-0509 , <http://dx.doi.org/10.1186/1752-0509-5-72>

This version is available at <https://strathprints.strath.ac.uk/35946/>

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (<https://strathprints.strath.ac.uk/>) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk

RESEARCH ARTICLE

Open Access

Exploring metabolic pathway disruption in the subchronic phencyclidine model of schizophrenia with the Generalized Singular Value Decomposition

Xiaolin Xiao¹, Neil Dawson^{2,3,4*}, Lynsey MacIntyre³, Brian J Morris^{2,5}, Judith A Pratt^{2,3,4}, David G Watson^{3,4} and Desmond J Higham^{1,4}

Abstract

Background: The quantification of experimentally-induced alterations in biological pathways remains a major challenge in systems biology. One example of this is the quantitative characterization of alterations in defined, established metabolic pathways from complex metabolomic data. At present, the disruption of a given metabolic pathway is inferred from metabolomic data by observing an alteration in the level of one or more individual metabolites present within that pathway. Not only is this approach open to subjectivity, as metabolites participate in multiple pathways, but it also ignores useful information available through the pairwise correlations between metabolites. This extra information may be incorporated using a higher-level approach that looks for alterations between a pair of correlation networks. In this way experimentally-induced alterations in metabolic pathways can be quantitatively defined by characterizing group differences in metabolite clustering. Taking this approach increases the objectivity of interpreting alterations in metabolic pathways from metabolomic data.

Results: We present and justify a new technique for comparing pairs of networks—in our case these networks are based on the same set of nodes and there are two distinct types of weighted edges. The algorithm is based on the Generalized Singular Value Decomposition (GSVD), which may be regarded as an extension of Principle Components Analysis to the case of two data sets. We show how the GSVD can be interpreted as a technique for reordering the two networks in order to reveal clusters that are exclusive to only one. Here we apply this algorithm to a new set of metabolomic data from the prefrontal cortex (PFC) of a translational model relevant to schizophrenia, rats treated subchronically with the N-methyl-D-Aspartic acid (NMDA) receptor antagonist phencyclidine (PCP). This provides us with a means to quantify which predefined metabolic pathways (Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolite pathway database) were altered in the PFC of PCP-treated rats. Several significant changes were discovered, notably: 1) neuroactive ligands active at glutamate and GABA receptors are disrupted in the PFC of PCP-treated animals, 2) glutamate dysfunction in these animals was not limited to compromised glutamatergic neurotransmission but also involves the disruption of metabolic pathways linked to glutamate; and 3) a specific series of purine reactions Xanthine \leftarrow Hypoxanthine \leftrightarrow Inosine \leftarrow IMP \rightarrow adenylosuccinate is also disrupted in the PFC of PCP-treated animals.

Conclusions: Network reordering via the GSVD provides a means to discover statistically validated differences in clustering between a pair of networks. In practice this analytical approach, when applied to metabolomic data, allows us to quantify the alterations in metabolic pathways between two experimental groups. With this new computational technique we identified metabolic pathway alterations that are consistent with known results. Furthermore, we discovered disruption in a novel series of purine reactions that may contribute to the PFC dysfunction and cognitive deficits seen in schizophrenia.

* Correspondence: N.Dawson@bio.gla.ac.uk

²Psychiatric Research Institute of Neuroscience in Glasgow (PsyRING),
Universities of Glasgow and Strathclyde, G12 8QQ, UK
Full list of author information is available at the end of the article

Background

Background in neuroscience and metabolomics

Schizophrenia is characterized by deficits in cognition known to be dependent upon the functional integrity of the prefrontal cortex (PFC). Furthermore, compromised PFC function in schizophrenia is supported by a multitude of neuroimaging studies reporting hypometabolism ('hypofrontality'), as evidenced by decreased blood flow or glucose utilization [1,2]. While the pathophysiological basis of PFC dysfunction in schizophrenia is not completely understood, a central role for NMDA receptor hypofunction is widely supported. For example, subchronic exposure to the NMDA receptor antagonist phencyclidine (PCP) induces cognitive deficits and a 'hypofrontality' which directly parallels that seen in schizophrenia [3-5]. Furthermore, subchronic PCP exposure induces alterations in GABAergic cell markers and 5-HT receptor expression in the PFC similar to those seen in this disorder [3,6,7]. While this evidence places NMDA receptor hypofunction central to the pathophysiology of PFC dysfunction in schizophrenia, the mechanisms through which NMDA hypofunction promotes PFC dysfunction are poorly understood.

Metabolomics is the comprehensive analysis of small molecule metabolites in biological systems [8]. It involves the study of the metabolome which is defined as all of the small molecular weight compounds within a sample that are required for metabolism, whose roles include growth and functionality [9-11]. Sample sources include bacteria, parasites, animals and humans and sample types can include biofluids, cells or tissue extracts. Metabolomics can be utilized as a tool for the characterization and quantification of all of the metabolites in a biological system. Its applications include profiling disease biomarkers [12,13], monitoring disease progression [14], investigating xenobiotic metabolism [15], investigating drug-induced toxicity [16,17] and investigating metabolism in genetically modified animals [18]. Mass spectrometry (MS) has been employed extensively as an analytical platform for metabolomics studies [19-21]. The popularity of this approach has increased over the last decade, in part due to the advent of high resolution Fourier transform mass spectrometers which offer improved reproducibility, accuracy and sensitivity. This makes mass spectrometry suitable for high throughput metabolomics studies [22]. In addition, the Orbitrap mass spectrometers that are now available offer similar performance to FT-MS systems without the need for a high strength magnetic field [23]. HILIC chromatography has been utilized as a separation technique prior to MS detection of polar metabolites in aqueous biofluids such as urine, serum and plasma [24-30].

Additionally, it has also been used for the detection of multiple neurotransmitters in primate cerebral cortex

[31]. HILIC chromatography has been chosen for metabolomic studies as it is useful for the analysis of highly polar metabolites which are poorly retained on reverse phase columns [32]. Detailed reviews of the principles and applications of HILIC have been previously outlined [25,33]. Here, HILIC-chromatography is utilized in combination with an LTQ-Orbitrap for metabolic profiling of metabolite extracts from the PFC of control and PCP-treated rats.

Metabolomics represents a robust approach through which alterations in diverse metabolic pathways may be determined at a biological systems level. In this way a metabolomics approach may prove useful in further elucidating the pathophysiological mechanisms contributing to PFC dysfunction in schizophrenia. Furthermore, this approach may also allow for the identification of PFC metabolic biomarkers for the cognitive deficits in this disorder. While the metabolomics approach can provide a rich and comprehensive set of data, the appropriate quantitative analysis of this data has not been adequately developed. In particular, the identification of statistical differences in metabolic pathways between experimental groups rather than the identification of statistical differences in individual metabolites alone represents a major challenge to quantitatively identifying metabolic alterations at a systems level from metabolomic data. One method through which statistical differences in metabolic pathways can be identified from metabolomic data involves the representation of this data as a large, complex network of nodes (single metabolites) connected by real-value edges (the correlation coefficient between two metabolites). This form of representation has high face validity as the relationship between two metabolites, in a given pathway, is governed by a single or series of enzymatic reactions that can be viewed as being represented by the correlation between the concentrations of the two metabolites. Another advantage is that metabolomic data consist of a range of metabolites detected in both of the experimental groups of interest meaning that these data can be expressed as two complex networks based upon the same set of nodes. This data structure is amenable to analysis through the application of the Generalized Singular Value Decomposition (GSVD) algorithm.

Background in network science and spectral methods

Large, complex interaction networks arise across many applications in science and technology [34-36]. Spectral methods, based on information computed from eigenvectors or singular vectors, have been used successfully to reveal fundamental network properties. For example, we may wish to cluster objects into groups [37], put objects into order [38] or discover specific patterns of

connectivity within subgroups [39-42]. In this work, we look at the case where two interaction data sets are available and the aim is discover differences between the two sets in the form of mutually exclusive clusters. For example, a given group of biologically defined entities, such as genes, proteins, metabolites or brain regions, may contain a subgroup that behaves in a coordinated manner under one condition, or in one organism, but not in another—the network with respect to one type of interaction contains a cluster that is not present in the other. We will show that the Generalized Singular Value Decomposition, which is becoming more widely used in computational biology [43,44] can be justified as the basis of a network reordering approach. We also consider how to quantify the statistical significance of network patterns that are uncovered.

Overall, this work develops and applies a novel algorithm in network science and shows that it reveals meaningful insights when applied to cutting-edge metabolomic data.

Results

Derivation of new algorithm

Suppose that the square, symmetric, real-valued matrices A and B in $\mathbb{R}^{N \times N}$ represent two different types of interaction between a set of N nodes. We have in mind the case where the weights play the role of correlation coefficients. Our aim is to discover clusters, in the sense of subsets of nodes that are mutually, pairwise, strongly connected through positive weights. The algorithm will also discover clusters of strong negative connectivity, although in practice this type of pattern is less likely to be present. However, we note that the arguments given below and the resulting algorithm remain valid in the case where the weights are non-negative, with zero representing the minimal level of similarity. The novelty of our approach is that in order to reveal interesting differences between the two types of connectivity data, we look for a set of nodes that form a good cluster with respect to A and a poor cluster with respect to B , or vice versa. As a starting point for a computational algorithm, we consider the identity

$$\|Ax\|_2^2 = \sum_{k=1}^N x_k^2 \deg_k^A + \sum_{i=1}^N \sum_{k=1}^N \sum_{l=1, l \neq k}^N a_{ik} a_{il} x_k x_l, \quad (1)$$

for $x \in \mathbb{R}^N$. Here $\|\cdot\|_2$ denotes the Euclidean norm and $\deg_k^A := \sum_{j=1}^N a_{kj}^2$ is one way to generalize the concept of out-degree to the case of a weighted network. Suppose we wish to split the nodes into two groups such that nodes within each group are well-connected but nodes across different groups are poorly connected. We could use an indicator vector $x \in \mathbb{R}^N$ to denote

such a partition, with $x_s = 1$ if node s is placed in group 1 and $x_s = -1$ if node s is placed in group 2.

Fixing on two nodes, k and l , we could argue that the existence of a third node, i , such that a_{ik} and a_{il} are both large and positive or both large and negative is evidence in favor of placing k and l in the same group (since they have in common a strong similarity or dissimilarity with node i). On the other hand small or oppositely signed values for a_{ik} and a_{il} is evidence in favor of placing k and l in different groups. In terms of the indicator vector, this translates to

1. $a_{ik} a_{il}$ large and positive \Rightarrow try to choose $x_k x_l = +1$,
2. $a_{ik} a_{il}$ small or negative \Rightarrow try to choose $x_k x_l = -1$.

Returning to the right-hand side of (1), we see that $\sum_{k=1}^N x_k^2 \deg_k^A$ is independent of the choice of indicator vector, and $\sum_{i=1}^N \sum_{k=1}^N \sum_{l=1, l \neq k}^N a_{ik} a_{il} x_k x_l$ gives a measure of how successfully we have incorporated the (possibly conflicting) desiderata in points 1 and 2 over all pairs k, l and third parties i . So we could judge the quality of an indicator vector by its ability to produce a large value of $\|Ax\|_2^2$, provided other constraints, such as balanced group sizes, were satisfied.

Analogously, we can argue that making $\sum_{i=1}^N \sum_{k=1}^N \sum_{l=1, l \neq k}^N a_{ik} a_{il} x_k x_l$ as negative as possible is a good way to avoid forming well-connected subgroups, and so the problem

$$\max_{x_s \in \pm 1, 1 \leq s \leq N} \frac{\|Ax\|_2^2}{\|Bx\|_2^2} \quad (2)$$

is a good basis for picking out strong clusters in A that are not present in B .

In general, optimizing over a large, discrete set of possibilities is computationally intractable, and hence we will follow the widely used practice of relaxing to an optimization over \mathbb{R}^N [37,45]. This approach goes back as far as the pioneering work of Fiedler [46] and has some theoretical underpinning in the case where a single network is analyzed [47,48]. So, instead of (2) we have

$$\max_{x \in \mathbb{R}^N, x \neq 0} \frac{\|Ax\|_2^2}{\|Bx\|_2^2}. \quad (3)$$

At this stage we recall that a general pair of matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times n}$ can be simultaneously factorized using the Generalized Singular Value Decomposition (GSVD) into

$$A = UCX^{-1} \quad \text{and} \quad B = VSX^{-1}, \quad (4)$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{p \times p}$ are both orthogonal, $X \in \mathbb{R}^{n \times n}$ is invertible, $C \in \mathbb{R}^{m \times n}$ and $S \in \mathbb{R}^{p \times n}$ are diagonal with nonnegative entries such that $C = \text{diag}(c_1, c_2, \dots$

c_n) and $S = \text{diag}(s_1, s_2, \dots, s_q)$ with $q = \min(p, n)$, and $0 \leq c_1 \leq c_2 \leq \dots \leq c_n$ and $s_1 \geq s_2 \geq \dots \geq s_q \geq 0$ [49]. The ratios $\lambda_i = c_i/s_i$ are the *generalized singular values* of A and B .

A key property of the GSVD is that the columns of X are stationary points of the function $f: \mathbb{R}^n \mapsto \mathbb{R}$ given by $f(x) = \|Ax\|_2 / \|Bx\|_2$, with the generalized singular values λ_i giving the corresponding stationary values. Hence, we may tackle the problem (3) through the GSVD. Columns 1, 2, 3, ... of X are candidates for finding good clusters in B that are poor clusters in A and, analogously, columns $N, N-1, N-2, \dots$ of X are candidates for finding good clusters in A that are poor clusters in B .

To transform back from real to discrete domains, we may use the ordering of the elements in x to define a new ordering for the two networks. More precisely, we relabel row and column i of A and B as row and column p_i where

$$p_i \leq p_j \Leftrightarrow x_i \leq x_j.$$

In this way, the existence or lack of clusters in each network becomes apparent from inspection of the heat map of the reordered matrix. This is the approach that we use. We will also show that p -values can be computed to quantify the statistical significance of the results. The issue of fully automating the choice of cluster size is left as future work.

A variant of the algorithm

In our context, the matrices A and B are square, with $m = n = p = N$. In this case, if we make the additional assumption that A and B are invertible it is known that the GSVD is closely related to the standard Singular Value Decompositions (SVD) of AB^{-1} and BA^{-1} . To see this, we could rearrange (4) into

$$AB^{-1} = UCS^{-1}V^T \quad \text{and} \quad BA^{-1} = VSC^{-1}U^T. \quad (5)$$

Alternatively, we may let $z = Ax$ or $y = Bx$ in (3), to obtain the quadratic problems

$$\max_{z \in \mathbb{R}^N, z \neq 0} \frac{\|z\|_2^2}{\|BA^{-1}z\|_2^2} \quad \text{or} \quad \max_{y \in \mathbb{R}^N, y \neq 0} \frac{\|AB^{-1}x\|_2^2}{\|y\|_2^2},$$

which can be solved through the standard SVD.

It is known from spectral graph theory that the dominant singular vectors give good directions in which to look for clusters [37,50]. Inverting the weight matrix reverses their importance (the singular value σ becomes σ^{-1}) and hence a spectral clustering approach applied to A^{-1} will typically find the opposite of good clusters—poorly connected nodes will be grouped together [51]. So, intuitively, forming AB^{-1} in (5) should produce a data matrix for which the SVD approach finds good

clusters for A and poor clusters for B . Analogously, the opposite holds for BA^{-1} .

Having interpreted the algorithm this way, it is then natural to consider the reverse products, $A^{-1}B$ and $B^{-1}A$, or, equivalently, to form the optimization problem

$$\max_{x \in \mathbb{R}^N, x \neq 0} \frac{\|B^{-1}x\|_2^2}{\|A^{-1}x\|_2^2}. \quad (6)$$

We may interpret (6) from the point of view that making $B^{-1}x$ large encourages poor clusters for B , while making $A^{-1}x$ small encourages good clusters for A . In this case, we would base our algorithm on the GSVD of A^{-1} and B^{-1} .

In the situation where A and B are both symmetric, corresponding to undirected networks, we have, from (4),

$$A^{-1} = (A^T)^{-1} = (X^{-T}CU^T)^{-1} = UC^{-1}X^T$$

and

$$B^{-1} = (B^T)^{-1} = (X^{-T}SV^T)^{-1} = VS^{-1}X^T.$$

Then we may appeal to the arguments given previously and use columns from the inverse of the third factor in the GSVD as the basis for reordering. With this approach we use columns of X^{-T} rather than columns of X . We emphasize that although this heuristic derivation used an assumption that A and B are invertible, the GSVD, and hence the final algorithm, applies in the non-invertible case. Also, the algorithms that we use do not require the computation of matrix inverses.

In tests on both synthetic and real network pairs, we found that this version of the algorithm was more effective, [52]. Hence, in this work we focus on the approach of reordering networks pairs via columns of X^{-T} . In summary, the first few columns of X^{-T} should give orderings that favor clusters in B rather than A and vice versa for the final few columns. In our computational examples, we used the `gsvd` routine built in to MATLAB <http://www.mathworks.com/>.

Synthetic test on binary networks

In this section we illustrate the algorithm in a simple, controlled case where we know the “correct” answer. We begin by considering binary networks, where results can be clearly visualized. We generated binary adjacency matrices A and B as shown in Figure 1. Here we have 20 nodes. In both networks, nodes 1-5 are well connected. In A there is a well connected cluster consisting of nodes 6-15, whereas in B there is a well connected cluster consisting of nodes 15-20. To make the test more realistic, the clusters are not perfect; there are both missing edges (false negatives) within the clusters

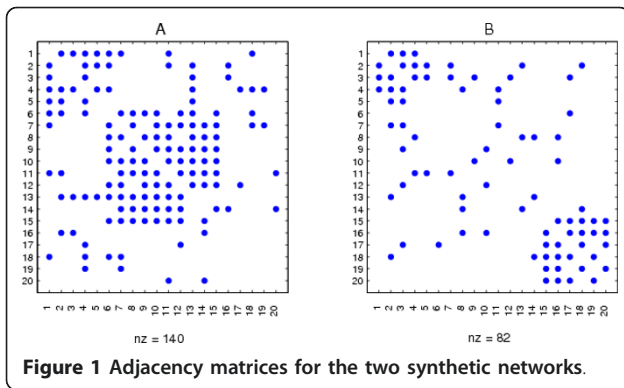


Figure 1 Adjacency matrices for the two synthetic networks.

and spurious edges (false positives) outside the clusters. Our aim is to test whether the algorithms can identify the clusters that are particular to each data set. We then show how statistical significance can be quantified.

We emphasize that the node labeling in Figure 1 was chosen purely to make the inherent structure visually apparent. Any spectral reordering algorithm should be invariant to a relabeling of the input data. In our context, this follows from the fact that for any permutation matrix P , the factorizations $A = UCX^{-1}$ and $B = V SX^{-1}$ are equivalent to $PAP^T = (PU)C(PX)^{-1}$ and $PBP^T = (PU)S(PX)^{-1}$. So, on the relabeled data matrices, (PX) plays the role that was played by X , and our algorithm reorders based on the appropriately permuted columns of X^{-T} , as required. In Figure 2 we show the same two data sets with an arbitrary relabeling in order to illustrate that the inherent structure is no longer apparent. In essence, we are hoping that the algorithm will find the structure that has been buried in Figure 2.

In Figure 3 we display the two adjacency matrices reordered with the algorithm; we show reordering with eight different columns of X^{-T} , four from each end of the spectrum. We see that mutually exclusive structures have been uncovered. The reordering from the first column begins with nodes 18, 20, 16, 15, 19, 17, which form a cluster in B , but not A . The final column begins

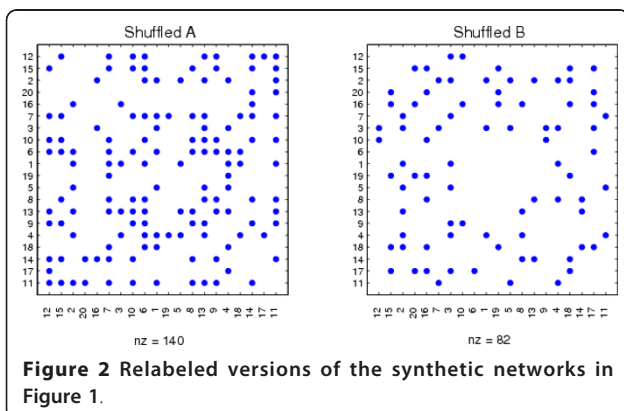


Figure 2 Relabeled versions of the synthetic networks in Figure 1.

by picking out nodes 7, 9, 10, 15, 14, 11, 6, 13, which form the bulk of the 6-15 cluster in A . Nodes 8 and 12, which are missing from this sequential ordering, are placed at the head of the ordering in the penultimate column, which begins 12, 8, 7, 10, 15, 14, 9, 11. So in summary, the 19th and 20th columns of X^{-T} each reveal almost complete information about the exclusive cluster in A , and between them they capture the full cluster.

Cluster validation

Suppose we find τ nodes giving a good cluster s for B but a poor cluster for A when the graphs are reordered by column ν from X^{-T} . Is this type of substructure likely to arise “by chance”? The following general approach can be used in order to determine a p -value, where we will regard a value below 0.05 as indicating statistical significance.

Initialization: Compute a measure of cluster quality, $c(A, B)$, for the promising substructure consisting of those τ nodes in networks A and B reordered by column ν .

Step 1: Randomize the networks and obtain new data sets \hat{A} and \hat{B} .

Step 2: Compute the GSVD for the randomized networks \hat{A} and \hat{B} and obtain a matrix \hat{X}^{-T} .

Step 3: Compute the measure $c(\hat{A}, \hat{B})$ for the τ node ‘cluster’ in \hat{A} and \hat{B} reordered by column ν from \hat{X}^{-T} .

p -value After performing M loops over Steps 1 to 3, compute a p -value as the proportion of $c(\hat{A}, \hat{B})$ samples that exceed $c(A, B)$.

For our cluster quality measure $c(A, B)$ we used

$$\frac{(\text{density of edges within the cluster in } B) / (\text{density of edges outside the cluster in } B)}{(\text{density of edges within the cluster in } A) / (\text{density of edges outside the cluster in } A)}$$

For these binary graphs, the density $f(s)$ of cluster s was defined as

$$f(s) = \frac{|E(s)|}{|s|}. \quad (7)$$

Here, $|E(s)|$ represents the actual number of edges in the object block s , and $|s|$ is the maximum possible number of edges.

For weighted graphs, in the case where the cluster is dominated by positive weights, we will generalize this to

$$f(s) = \frac{w(s)}{|s|}. \quad (8)$$

Here, $w(s)$ denotes the average weight in block s . We note the denominator $|s|$ cancels when ratios are computed in the p -value algorithm.

In Figure 3, we see that eight nodes 7,9,10,15,14,11,6,13 form a cluster in A , but not in B ,

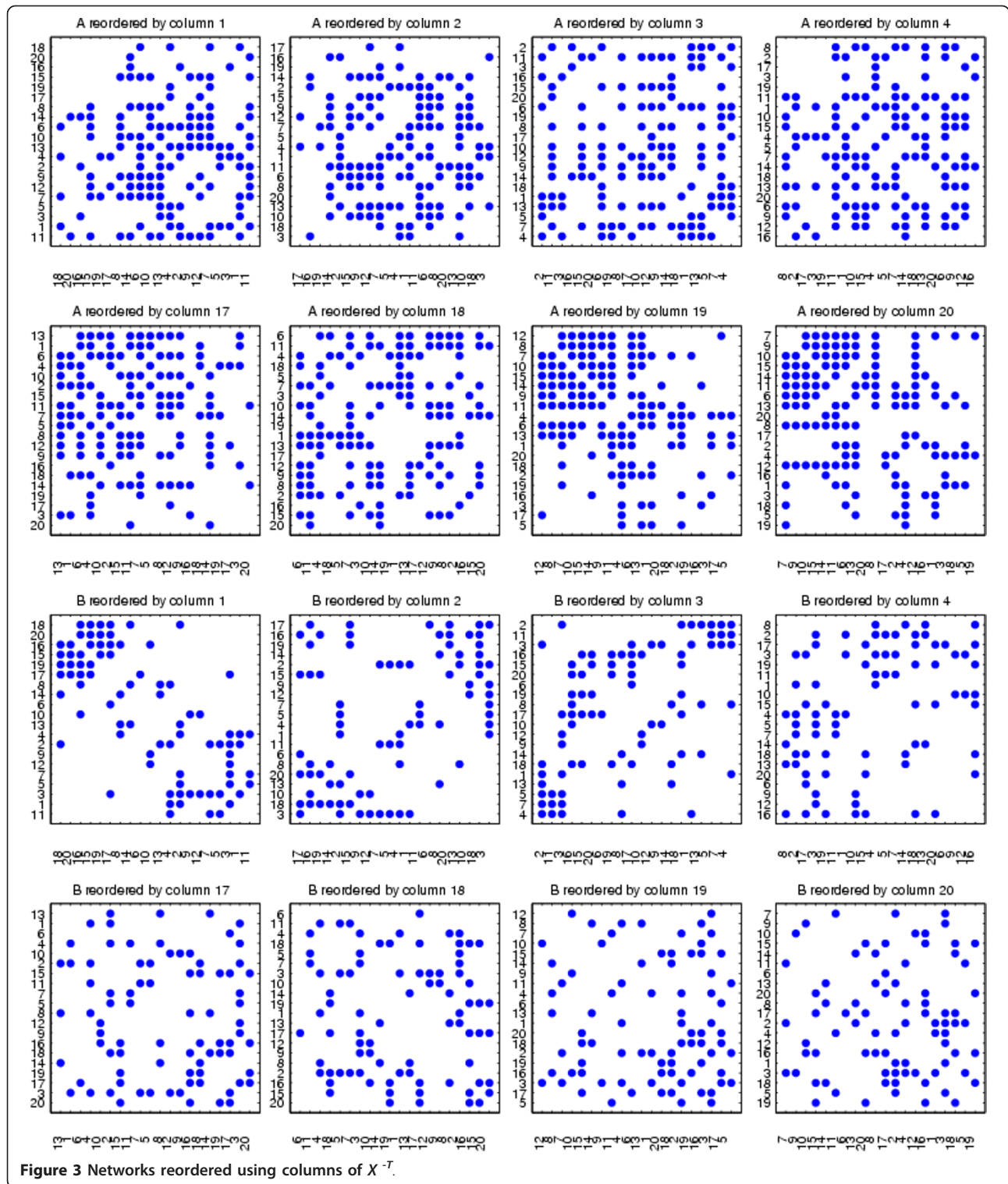


Figure 3 Networks reordered using columns of X^{-T} .

when the synthetic data is reordered with the final column of X^{-T} . Applying the procedure above, using permutation to randomize the networks $M = 1000$ times as described below, we obtained a p -value of 0.007. Applying the same procedure, we also obtained a p -value of

0.029 for the first 6 nodes 18, 20, 16, 15, 19, 17 when the synthetic data is reordered with the first column of X^{-T} , which visually form a cluster in B , but not in A . These p -values (< 0.05) both indicate that the results are statistically significant. As a further test, we arbitrarily

selected the subnetworks of A and B composed of nodes 2,4,12,16,1,3,18, which correspond to the 12th to 18th components of the sorted final column from X^{-T} . In this case, we would not expect to find a significant result. This is reflected in the large p -value of 0.844. In more exhaustive experiments, three randomization methods were tested [52]:

- **Erdős-Rényi:** generate a classical random graph with the appropriate number of edges.
- **Redistribution:** redistribute the entries in each row and each column of A , and perform the same operations on B .
- **Permutation:** reorder the nodes in A and B and choose the first τ nodes in this new ordering. In this case, recomputation of the GSVD in Step 2 is not necessary, due to the permutation invariance of the factorization.

Of those three approaches, Erdős-Rényi may be the most commonly used method to randomize a binary network, whereas permutation extends most naturally to the case of weighted edges, so we used permutation in the test shown here. We also tested another simple cluster quality measure which is the ratio of density of edges within the cluster in one graph and that in the other graph.

These variations were studied within this general methodology on both real and synthetic data sets [52]. In all cases, comparable p -values were produced.

Synthetic test on correlation networks

Having tested the algorithm on binary networks, we now consider the case where weighted edges arise as correlation coefficients.

First, we generate two correlation matrices A and B as shown in Figure 4. Here, each graph has 20 nodes, and each entry is real valued, representing the correlation coefficient between the corresponding nodes. The same cluster patterns given for the synthetic binary matrices in Figure 1 were built in to the synthetic correlation data: nodes 1-5 are well connected in both networks; in

A there is a well connected cluster consisting of nodes 6-15, whereas in B there is a well connected cluster consisting of nodes 15-20. Some noise was added to the clusters to make this test more realistic.

More precisely, in our computation, the value of each entry (the correlation coefficient) in A and B as shown in Figure 4 is generated from a pair of 20×50 rectangular matrices D_a and D_b . The corresponding cluster patterns are built from signals. Figure 5 shows the nine signals that take part in the data. These are row vectors with 50 elements. We use $v^{[1]}, v^{[2]}, v^{[3]}, \dots, v^{[9]}$ to denote them. From these signals, we set up two matrices

- $D_a \in \mathbb{R}^{20 \times 50}$: the first 5 rows are linear combinations of $v^{[1]}, v^{[2]}, v^{[3]}, v^{[4]}, v^{[5]}, v^{[6]}$ and $v^{[7]}$. Rows 6 to 15 are combinations of $v^{[7]}$ and $v^{[8]}$. The remaining rows (rows 16 to 20) are Gaussian pseudorandom numbers.
- $D_b \in \mathbb{R}^{20 \times 50}$: the first 5 rows are linear combinations of $v^{[1]}, v^{[2]}, v^{[3]}, v^{[4]}, v^{[5]}, v^{[6]}$ and $v^{[7]}$. Rows 6 to 14 are Gaussian pseudorandom numbers. The remaining rows (rows 15 to 20) are combinations of $v^{[4]}$ and $v^{[9]}$.

Building up the rows from the underlying signals in this manner allowed us to construct the correlation patterns seen in Figure 4.

Although the algorithm is invariant to permutation, for visual clarity, we also shuffled the synthetic correlation data sets A and B before applying our algorithm to them. Figure 6 shows the same synthetic correlation data sets with an arbitrary relabeling.

We present the results from our algorithm in Figures 7 and 8. We show the relabeled A and B reordered with two extreme columns of X^{-T} , one from each of the two ends of the spectrum. The reorderings reveal the mutually exclusive cluster structures of A and B . We also applied the cluster validation method to the structures uncovered by the reorderings using random

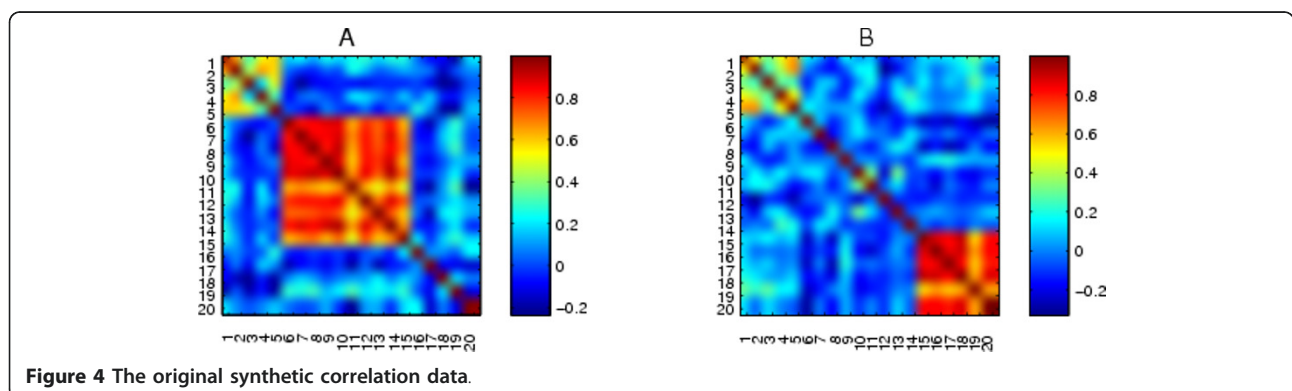
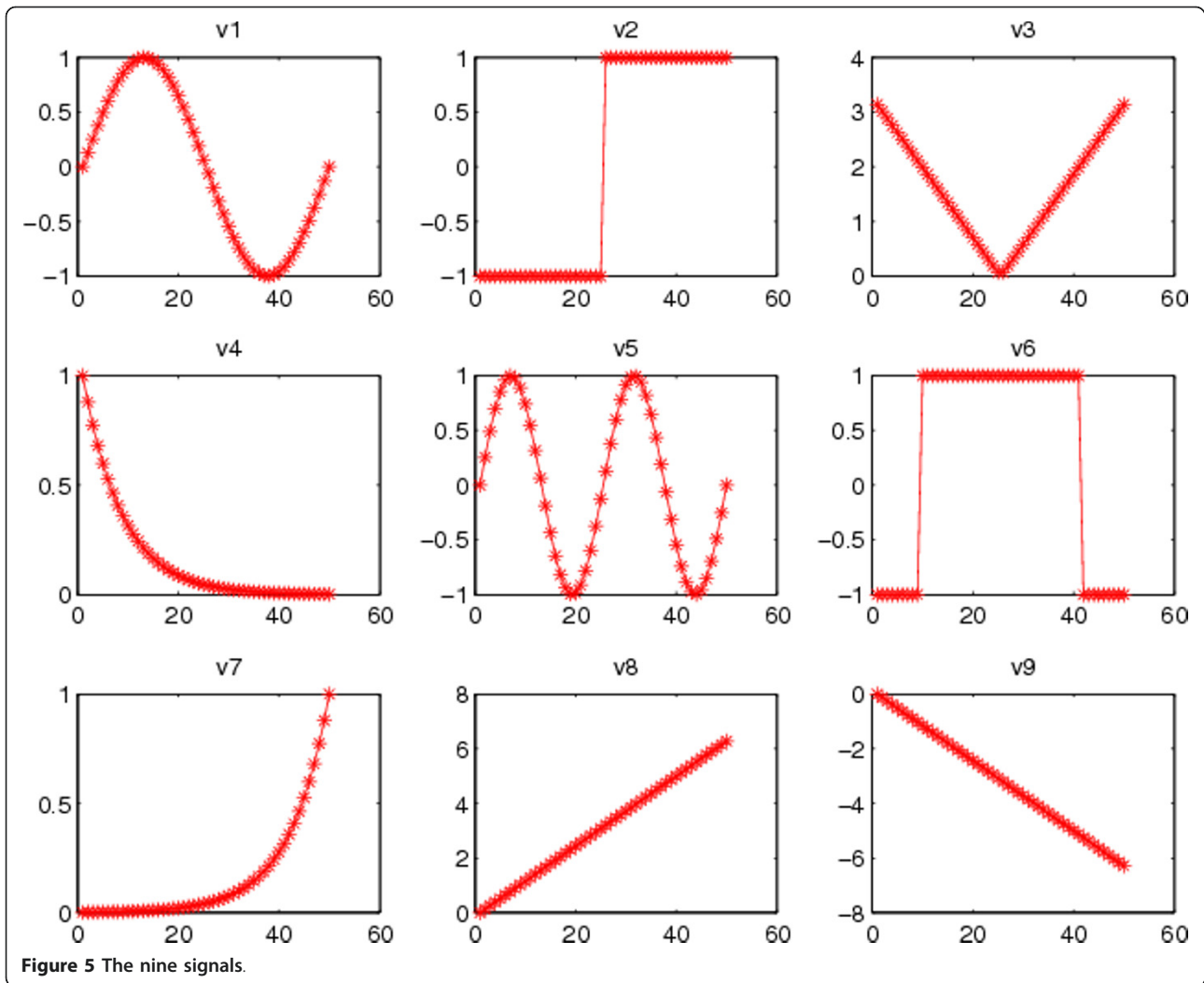


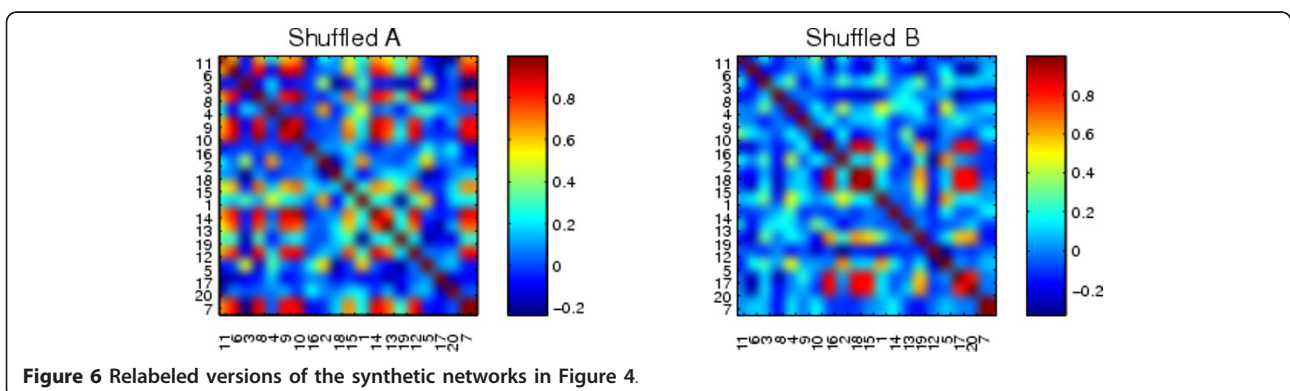
Figure 4 The original synthetic correlation data.



permutation. In Figure 7 we see that the first column of X^{-T} picks out the continuous nodes 17, 15, 20, 18, 16, 19, which form a good cluster in B but not in A ($p < 0.001$). The reordering from the final column of X^{-T} shown in Figure 8 reveals that the 6-15 cluster in A but

not in B was completely uncovered by the nodes 10, 14, 12, 9, 8, 7, 6, 13, 15, 11 (at the top left hand side of the heatmaps, $p < 0.001$).

In summary, this additional synthetic test illustrates that our GSVD based algorithm can be extended to



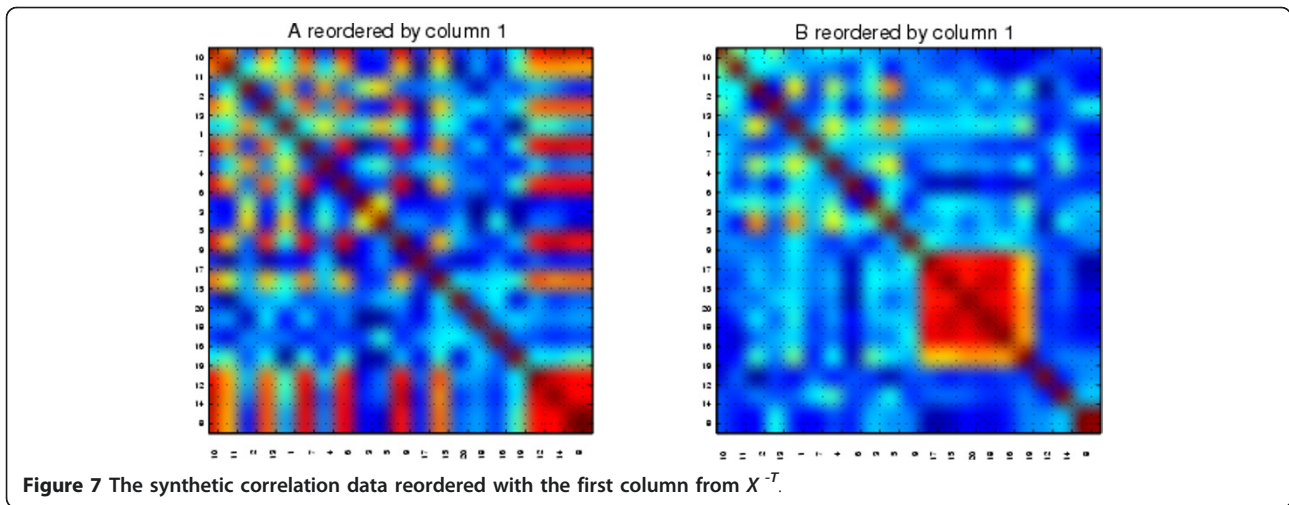


Figure 7 The synthetic correlation data reordered with the first column from X^{-T} .

reveal the pattern difference between two relative correlation matrices in terms of clustering.

Quantitative determination of metabolic pathways disrupted in the prefrontal cortex of PCP-treated animals
 SIEVE analysis (Thermo-Fisher Scientific) revealed significant PCP-induced alterations in the level of specific metabolites in the PFC of PCP-treated rats (Table 1 Additional File 1). These changes were evident in multiple metabolic pathways as defined by the Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolite pathways database. Significant changes were evident in (i) glutamate metabolism (3 metabolites [m, n]), (ii) the alanine, aspartate and glutamate pathway (2 metabolites [n]), (iii) phenylalanine, tyrosine and tryptophan metabolism (3 metabolites [a]), (iv) purine metabolism (2 metabolites [o]) and (v) butanoate metabolism (2 metabolites [k]). This suggested that these metabolic pathways are disrupted in the PFC of PCP-treated animals.

However, this simple level of analysis prevents any quantitative and statistically rigorous determination of the predefined (KEGG) metabolic pathways disrupted in the PFC of PCP-treated animals.

In the context of this study the aim of applying the GSVD algorithm to metabolomic data from control and PCP-treated animals was to quantitatively determine which predefined metabolic pathways were altered in PCP-treated animals. The inter-metabolite Pearson's correlation coefficient (partial correlation) was used as the metric of the functional association between each pair of metabolites and was generated from the metabolite peak intensities, as determined by Liquid Chromatography Mass Spectrometry (LC-MS), across all animals within the same experimental group (i.e. either control or PCP-treated). These correlations were Fisher transformed to give the correlation data a normal distribution. This resulted in a pair of symmetric, square, real-valued $\{98 \times 98\}$ partial correlation matrices (Control

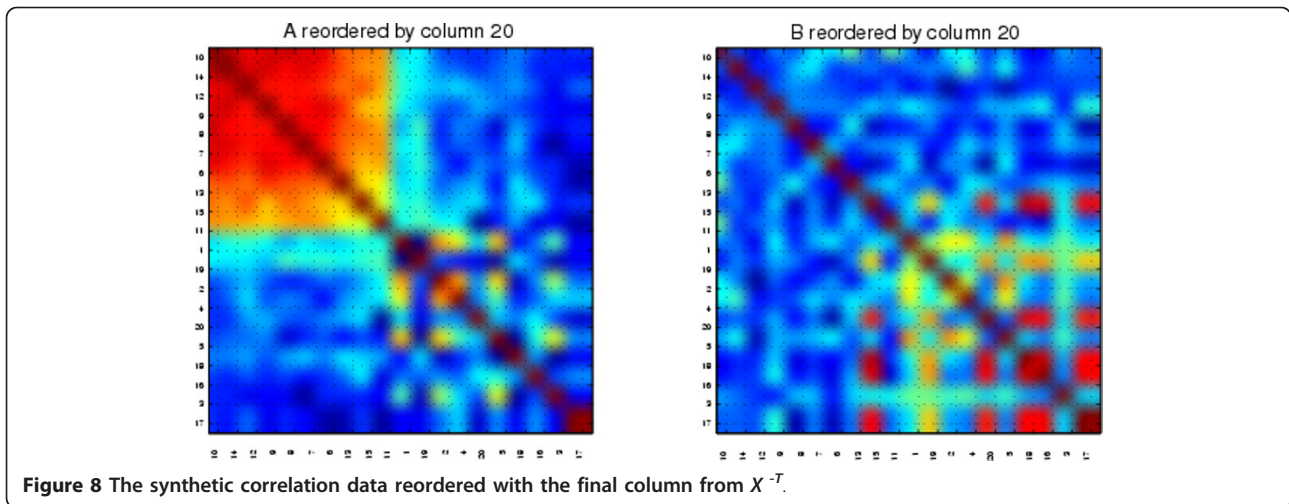


Figure 8 The synthetic correlation data reordered with the final column from X^{-T} .

Table 1 PCP-induced alterations in PFC metabolite levels as determined by SIEVE analysis

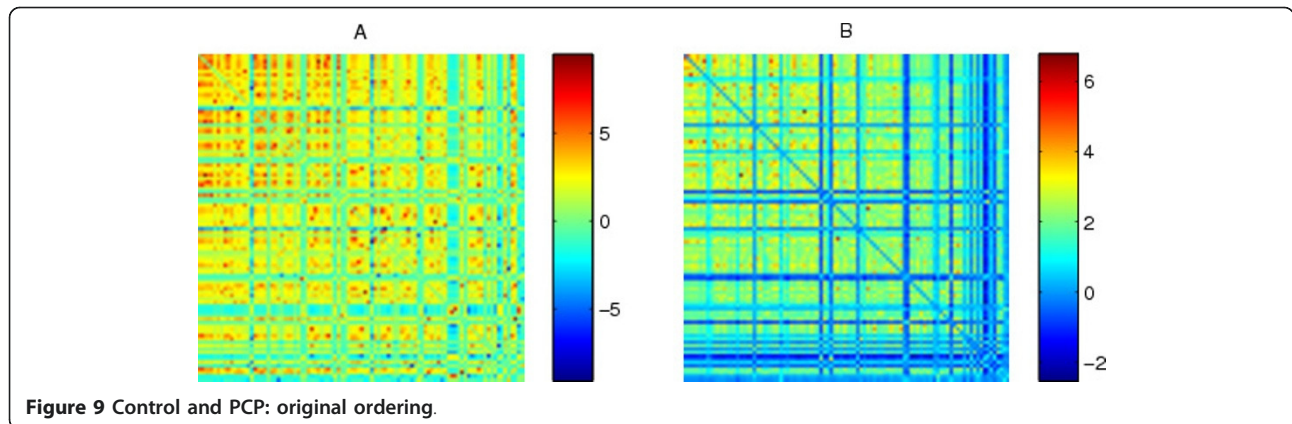
Formula	Metabolite	Metaboite KEGG ID	KEGG Pathways	P-value	Ratio
C ₉ H ₁₁ NO ₃	L-Tyrosine	c00082	ko00350, ko00360, ko00400	0.001	0.584
C ₁₀ H ₁₇ N ₃ O ₆	gamma Glutamylglutamine	NA	NA	0.007	0.673
C ₆ H ₁₃ N ₃ O ₃	L-Citrulline	c00327	ko00330	0.007	0.709
C ₃ H ₇ NO ₂ S	L-Cysteine	c00097	ko00260, ko00270, ko00430, ko00480, ko00730, ko00770, ko00920	0.012	0.445
C ₈ H ₉ NO	2-Phenylacetamide	c02505	ko00360	0.015	0.561
C ₉ H ₈ O ₃	Phenylpyruvate	c00166	ko00360, ko00400	0.016	0.57
C ₄ H ₆ O ₂	2,3-Butanedione	c00741	map00650	0.017	0.786
C ₄ H ₅ N ₃ O	Cytosine	c00380	ko00240	0.019	0.665
C ₀₄ H ₉ NO ₂	GABA	c00334	ko00250, ko00330, ko00410, ap00650, ko04080	0.021	0.804
C ₉ H ₁₇ NO ₄	O-Acetylcarnitine	c02571	ko00250	0.022	2.649
C ₁₄ H ₁₈ N ₅ O ₁₁ P	Adenylosuccinate	c03794	ko00230, ko00250	0.029	3.276
C ₅ H ₅ N ₅ O	Guanine	c00242	ko00230	0.035	0.593
C ₇ H ₁₆ NO ₃	Carnitine	c00487	ko00310	0.037	0.819

Table 1 shows the molecular formula, tentative metabolite identity and the KEGG pathways in which a metabolite is involved. Only metabolites found to be significantly different between the two experimental groups by SIEVE analysis (see Methods section) are shown. Full data for all metabolites detected in the PFC of control and PCP-treated rats are shown in Table S1 (Additional File 1). The most prominent alterations in KEGG defined metabolic pathways appeared to be in (i) alanine, aspartate and glutamate metabolism (3 metabolites [ko00250]), (ii) phenylalanine, tyrosine and tryptophan metabolism (3 metabolites [ko00360]), (iii) purine metabolism (2 metabolites [ko00230]) and (iv) butanoate metabolism (2 metabolites [ko00650]). KEGG defined metabolic pathways; ko00250: Alanine, Aspartate and Glutamate metabolism; ko00330: Arginine and Proline metabolism; ko00410: beta-Alanine metabolism; map00650: Butanoate metabolism; ko00270: Cysteine and Methionine metabolism; ko0480: Glutathione metabolism; ko00260: Glycine, Serine and Threonine metabolism; ko00430: Methionine metabolism; ko04080: Neuroactive ligand-receptor interaction; ko00770: Pantoate and CoA biosynthesis; ko00360: Phenylalanine metabolism; ko00400: Phenylalanine, Tyrosine and Tryptophan biosynthesis; ko00230: Purine metabolism; ko00240: Pyrimidine metabolism; ko00920: Sulphur metabolism; ko00430: Taurine and Hypotaurine metabolism; ko00730: Thiamine metabolism; ko00350: Tyrosine metabolism; ko00400: Tyrosine and Tryptophan biosynthesis. NA denotes a metabolite not associated with a KEGG compound ID or KEGG pathway.

animals: Additional File 2 PCP-treated animals: Additional File 3). Each within-group matrix represents the specific association strength between each of the 9506 possible pairs of metabolites in that experimental group. In the simplest biological case the correlation coefficient between two metabolites (nodes) in the matrix represents the series of enzymatic reactions responsible for converting one metabolite into another. However, it should be noted that this simple interpretation does not account for the complex relationships that may influence the correlation between two metabolites, such as the involvement of metabolites in alternative, often parallel, metabolic pathways. There are important limitations that must be recognized when modeling metabolomic data as a complex network of interactions between metabolites (as defined by the correlation that exists between them) such as the potential for correlations to exist between metabolites that are not biologically relevant. The impact of such erroneous associations on the interpretation of the data as outlined in this paper will be limited by the approach of characterizing alterations at the level of metabolic pathways, involving multiple metabolites (the approach taken in this study), rather than considering the disruption of single correlation coefficient between two metabolites.

Our network treats interactions between molecules as bidirectional, and so the set of interactions between

molecules forms an undirected weighted network. In essence the GSVD algorithm allows the reordering of the two experimental matrices *A* (control animals) and *B* (PCP-treated animals) with the aim of discovering a new node (metabolite) ordering that reveals clusters of nodes that exhibit strong connectivity (mutual weights) in one network but not the other. In the context of this data the GSVD algorithm was used to identify clusters of metabolites present in one experimental group that are not present in the other with the aim of identifying those metabolic pathways in the PFC disrupted by PCP treatment. Once the matrices had been reordered through the GSVD algorithm the significant presence of a cluster in the given network was statistically tested by comparison of the cluster quality measure in the real networks relative to that in 1000 random permutations of the initial matrices. The original metabolomic networks are shown in Figure 9, where matrix *A* represents control animals and *B* represents PCP-treated animals. Figures 10 and 11 show the networks reordered by the first and the final column of X^{-T} , respectively. The original position of each metabolite detected by LC-MS (Figure 9) and its new position in each of the reordered matrices (Figures 10 and 11) are shown in Additional File 4. Visually, in Figure 10 there was no obvious pattern of clustering that would identify significant clusters of metabolites present in PCP-treated animals that were

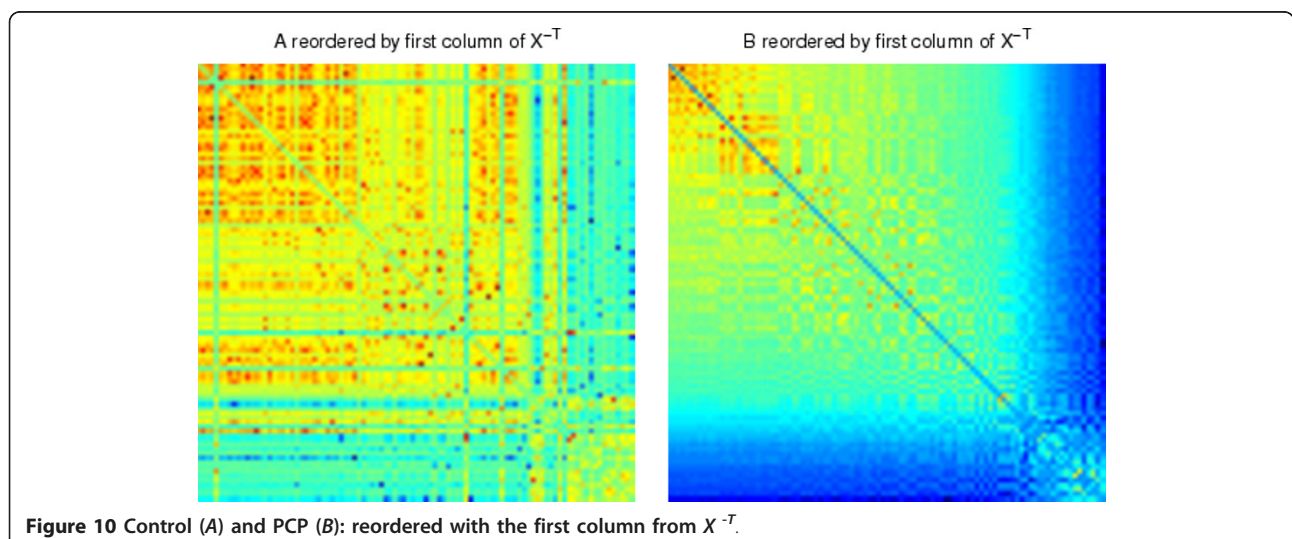


not present in controls. In contrast, in Figure 11 there appeared to be clusters of metabolites present in the PFC of control animals that were not present in PCP-treated animals (top left and bottom right hand side of the heatmap). For Figure 11 the significance of the top cluster (first 22 nodes in the reordering, $p < 0.001$) and the bottom cluster (last 18 nodes in the reordering, $p < 0.001$) was confirmed, indicating that there were clusters of metabolites significantly present in control (A) animals that were not present in PCP-treated (B) animals. The identity of the metabolites, the KEGG pathways in which each metabolite is involved, and the PCP-induced alteration in the overt level of each metabolite (as determined by SIEVE analysis) are shown in Tables 2 and 3 for the top and bottom cluster, respectively. In contrast to the metabolite clustering shown in Figure 11 there was no evidence in Figure 10 for any significant cluster of metabolites present in PCP-treated animals (B) that was not present in control (A) animals: (i) potential top cluster [first 10 nodes] $p = 0.421$; (ii) potential middle cluster [nodes 18-25] $p = 0.494$.

Rigorous significance testing, involving multiple potential metabolite clusters, confirmed that there were no significant clusters of metabolites in PCP-treated animals that were not present in controls (Figure 10). Following significance testing of potential metabolite clusters in the GSVD reordered matrices, hypergeometric probability (described in the Methods section) was applied to test the significance of KEGG defined metabolite pathway over-representation in these clusters. The results for hypergeometric probability testing are shown in Tables 4 and 5.

Discussion

Through its application to metabolomic data we have clearly demonstrated the added value that can be gained from applying the GSVD algorithm to two sets of complex, network data based upon the same set of nodes. In particular, we have demonstrated that the combined application of the GSVD algorithm with hypergeometric probability analysis provides an analytical framework by which statistical alterations in predefined metabolic



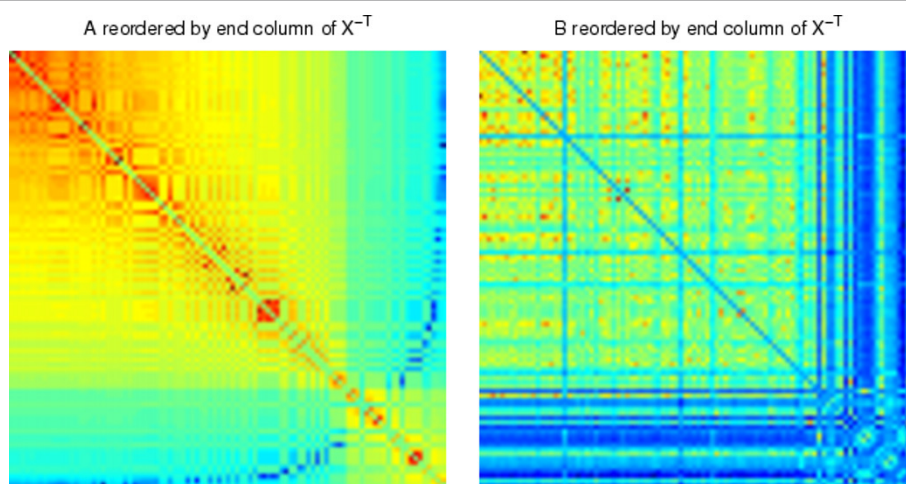


Figure 11 Control (A) and PCP (B): reordered with the final column of X^{-T} .

Table 2 Metabolite identities and their relevant KEGG pathways in the top cluster of Figure 11

Formula	Metabolite	Metaboite KEGG ID	KEGG Pathways	P -value	Ratio
$C_5H_{10}N_2O_3$	L-Glutamine	c00064	Ko00230, ko00240, ko00250, ko00330	0.522	0.959
H_3PO_4	Phosphoric acid	c00009	ko00190	0.254	0.915
$C_5H_7NO_3$	1-Pyrroline-4-hydroxy-2-carboxylate	c04282	ko00330	0.781	0.981
$C_4H_9N_3O_2$	Creatine	c00300	ko00330, ko00260	0.551	0.953
$C_4H_9NO_2$	GABA	c00334	ko00250, ko00330, ko00410, ko04080, map00650	0.021	0.804
$C_4H_7NO_4$	L-Aspartate	c00049	ko00250, ko00260, ko00270, map00300, ko00330, ko00340, ko00410, ko00760, ko00770, ko04080	0.319	0.916
$C_4H_7NO_2$	1-Aminocyclopropane-1-carboxylate	c01234	ko00270, ko00640	0.590	0.951
$C_5H_5N_5O$	Guanine	c00242	ko00230	0.035	0.593
$C_5H_9NO_4$	Glutamate	c00025	ko00250, ko00330, ko00340, ko00471, ko04080, ko00480, map00650	0.845	0.985
C_4H_7NO	Hydroxymethylpropanitrile	NA	NA	0.098	0.842
$C_6H_6N_2O$	Nicotinamide	c00153	ko00760	0.440	0.917
$C_4H_6O_2$	2,3-Butanedione	c00741	map00650	0.017	0.786
$C_6H_{12}O_4$	Pantoate	c00552	ko00770	0.722	0.963
$C_{15}H_{23}N_5O_{14}P_2$	ADP-ribose	c00301	ko00230	0.058	677.029
$C_3H_7NO_3$	L-Serine	c00065	ko00260, ko00270, ko00600, ko00920, ko00680	0.316	0.856
$C_4H_5N_3O$	Cytosine	c00380	ko00240	0.019	0.665
$C_2H_7NO_3S$	Taurine	c00245	ko00430, ko04080	0.936	0.995
$C_4H_5NO_3$	Maleamate	c01596	ko00760	0.372	0.927
$C_2H_8NO_4P$	Ethanolamine phosphate	c00346	ko00260, ko00564, ko00600	0.373	0.889
	Unknown ID	NA	NA	0.271	1.395
$C_5H_{11}NO_3$	Hydroxyvaline	NA	NA	0.585	0.946
$C_6H_{13}N_3O_3$	L-Citrulline	c00327	ko00330	0.007	0.709

Table 2 shows the top cluster of metabolites identified by the GSVD algorithm that are present in the PFC of control but not PCP-treated animals (Figure 11). The molecular formula, tentative molecular identity, its KEGG compound identity and the KEGG metabolic pathways in which a given metabolite is involved are also shown. The key for each KEGG pathway identity is shown in Table 4. The p -values and ratio change reported for each metabolite in this table are those calculated by SIEVE analysis. Those metabolites found to be significantly different between the two groups by analysis are highlighted in bold. While SIEVE analysis fails to attribute significance ($p < 0.05$) to PCP-induced alterations in the overt concentration of many metabolites in this cluster, GSVD analysis reveals that the relationship between these metabolites is significantly altered by PCP treatment ($p < 0.001$), highlighting the specific metabolic pathways that may be disrupted in the PFC of PCP-treated animals. The most prominent alterations in KEGG defined pathways in this cluster were in (i) Arginine and Proline metabolism (7 metabolites [ko00330]) (ii) Glycine, Serine and Threonine metabolism (3 metabolites [ko00260]) and (iii) KEGG defined neuroactive ligands (4 metabolites [ko04080]).

Table 3 Metabolite identities and their relevant KEGG pathways in the bottom cluster of Figure 11

Formula	Metabolite	Metaboite KEGG ID	KEGG Pathways	P -value	Ratio
$C_5H_4N_4O_2$	Xanthine	c00385	ko00230	0.339	0.508
$C_{10}H_{16}N_2O_7$	Gamma-glutamylglutamic acid	NA	NA	0.143	0.54
$C_{14}H_{26}O_2$	Myristoleic acid	c08322	NA	0.689	0.623
$C_5H_4N_4O$	Hypoxanthine	c00262	ko00230	0.115	0.569
$C_{17}H_{37}NO_2$	Heptadecaspinganine	NA	NA	0.733	0.769
$C_{10}H_{13}N_4O_8P$	Inosine monophosphate	c00130	ko00230	0.461	0.73
$C_{10}H_{17}N_3O_6$	Peptide fragment (Arg-Arg-Gln)	NA	NA	0.775	1.183
$C_6H_{15}NO_3$	Triethanolamine	c06771	ko00564	0.691	1.207
$C_9H_{14}N_4O_3$	Carnosine	c00386	ko00340, ko00410	0.872	1.128
$C_{10}H_{12}N_4O_5$	Inosine	c00294	ko00230	0.090	0.6
$C_{15}H_{12}O_5$	Narigenin	c00509	NA	0.196	0.862
$C_{10}H_{17}N_3O_6$	gamma Glutamylglutamine	NA	NA	0.007	0.673
$C_{26}H_{42}N_7O_{20}P_3S$	2-Hydroxyglutaryl-CoA	c03058	map00650	0.179	0.715
$C_{31}H_{54}N_7O_{17}P_3S$	Decanoyl-CoA	c05274	ko00071	0.410	1.312
$C_{25}H_{44}NO_7P$	2- Aminoethylphosphocholate	c05683	ko00440	0.243	0.662
$C_{22}H_{26}O_6$	Eudesmin	NA	NA	0.084	0.493
$C_3H_7NO_2S$	L-Cysteine	c00097	ko00260, ko00270, ko00430, ko00480, ko00730, ko00770, map00920	0.012	0.445
$C_3H_7O_6P$	Glycerone phosphate	c00111	ko00010, ko00051, ko00052, ko00561, ko00562, ko00564, ko00620	0.063	0.381

Table 3 concerns the bottom cluster of metabolites identified by the GSVD algorithm that are present in the PFC of control animals but not PCP-treated animals. The molecular formula, tentative molecular identity, its KEGG compound identity and the KEGG metabolic pathways in which a given metabolite is involved are shown. The identity of each KEGG pathway ID is shown in Table 5. The *p*-values and ratio change reported for each metabolite in this table are those calculated by SIEVE analysis. Those metabolites found to be significantly different between the two groups by analysis are highlighted in bold. While SIEVE analysis fails to attribute significance ($p < 0.05$) to PCP-induced alterations in the overt concentration of many metabolites for many metabolites in this cluster, the PCP/Control ratio suggests that the levels of many of these metabolites are markedly altered by PCP-treatment. GSVD analysis reveals that the relationship between the levels of these metabolites in this cluster are significantly altered by PCP-treatment ($p < 0.001$) highlighting specific metabolic pathways that may be disrupted in the PFC of PCP-treated animals. There appears to be an overabundance of Purine (4 metabolites [ko00230]) and Glycerophospholipid (2 metabolites [ko00564]) in the bottom cluster.

pathways between experimental groups can be defined from complex metabolomic data. There is a great unmet need for this type of analytical approach in metabolomics, as well as in the other -omics fields (e.g. transcriptomics), which allows the quantification of alterations at the biological systems (pathways) level rather than simply identifying significant alterations of discrete measures (i.e. single metabolites).

Through the application of this analytical approach we identified statistically significant alterations in specific, pre-defined metabolic pathways (KEGG database pathways) that may contribute to PFC dysfunction in PCP-treated animals, and so in schizophrenia. This included the disruption of the (1) Alanine, Aspartate and Glutamate [ko00250], (2) Arginine and Proline [ko00330], (3) Butanoate [ko00650], (4) Nicotinate and Nicotinamide [ko00760], (5) Glycine, Serine and Threonine metabolic pathways as well as an imbalance in (6) metabolites active as neurotransmitter ligands [ko04080]. The disruption of metabolic pathways involving glutamate in the PFC of PCP-treated rats seems particularly pertinent given the reported alterations in extracellular glutamate

availability in the PFC following repeated PCP treatment [53] and the central hypothesis of hypofunctional glutamatergic PFC neurotransmission in schizophrenia [54,55]. In addition to altered glutamate metabolism there was also evidence to support an imbalance in multiple metabolites known to be active at glutamate receptors. This included an imbalance in the relationship between glutamate, L-aspartate and Taurine (Table 2) which are all known to be active at glutamate receptors. Furthermore, evidence for the disruption of glycine, serine and threonine metabolism may suggest that glycine and serine activity as co-agonists at the NMDA receptors may be disrupted in the PFC of PCP-treated animals. However, it is important to note that we failed to detect glycine levels in our samples and serine levels appear to be overtly unchanged. The possibility of altered glycine levels in the PFC of PCP-treated rats warrants further investigation given the ability of glycine and NMDA receptor glycine site agonists to reverse subchronic PCP-induced alterations in PFC dopaminergic neurotransmission [56,57], which may be central to the impact of subchronic PCP treatment on cognition.

Table 4 Hypergeometric probability of KEGG defined metabolic pathways in the top cluster in Figure 11

KEGG Path-way Identity	KEGG Pathway	Number of metabolites in cluster(A)	Total number of pathway metabolites detected (B)	Hypergeometric Probability ($P(X) \geq k$)
ko00190	Oxidative phosphorylation	1	1	0.224
ko00230	Purine metabolism	3	13	0.598
ko00240	Pyrimidine metabolism	2	6	0.406
ko00250	Alanine, Aspartate and Glutamate metabolism	4	7	0.043
ko00260	Glycine, Serine and Threonine metabolism	4	7	0.043
ko00270	Cysteine and Methionine metabolism	3	7	0.186
map00300	Lysine biosynthesis	1	3	0.538
ko00330	Arginine and Proline metabolism	7	10	0.001
ko00340	Histidine metabolism	2	5	0.312
ko00410	beta-Alanine metabolism	2	5	0.312
ko00430	Taurine and Hypotaurine metabolism	1	3	0.538
ko00471	D-glutamine and D-glutamate metabolism	1	1	0.224
ko00480	Glutathione metabolism	1	5	0.728
ko00564	Glycerophospholipid metabolism	1	11	0.949
ko00600	Sphingolipid metabolism	2	3	0.126
ko00640	Propanoate metabolism	1	2	0.400
map00650	Butanoate metabolism	3	4	0.034
ko00680	Methane metabolism	1	1	0.224
ko00760	Nicotinate and Nicotinamide metabolism	3	4	0.034
ko00770	Pantothenate and CoA biosynthesis	2	5	0.312
ko00920	Sulphur metabolism	1	3	0.538
ko04080	Neuroactive ligand-receptor interaction	4	7	0.043

Table 4 shows the hypergeometric probability of at least the observed number of metabolites arising by chance for a given KEGG defined metabolic pathway in the top cluster of Figure 11, identified through the GSVD algorithm as being present in control but not PCP-treated animals. Further computational details are given in the Methods section. The cluster size was 22 metabolites from a total population of 98. There was a significant over representation of metabolites of (i) Alanine, Aspartate and Glutamate metabolism [ko00250], (ii) Arginine and Proline metabolism [ko00330], (iii) Butanoate metabolism [ko00650], (iv) Nicotinate and Nicotinamide metabolism [ko00760], (v) Glycine, Serine and Threonine metabolism and (vi) those metabolites active as neurotransmitter ligands [ko04080] (all highlighted in bold) suggesting that these pathways are disrupted in the PFC of PCP-treated animals.

Altered glycine, serine and threonine metabolism in the PFC of PCP-treated animals is also consistent with the hypothesis that glycine and serine represent potential therapeutic targets for the treatment of schizophrenia [58]. In addition, we found evidence to suggest that GABA neurotransmission was also significantly decreased in the PFC of PCP-treated rats, which may relate to the compromised integrity of GABAergic interneurons in these animals [3,6], which closely resemble the GABAergic interneuron alterations seen in schizophrenia. The imbalance in glutamate, glutamine and GABA levels identified in the PFC of PCP-treated rats may directly contribute to the hypofrontality (glucose hypometabolism) seen in these animals, as detected

using the ^{14}C -2-deoxyglucose imaging technique [4], as all of these metabolites are intimately linked through metabolic pathways and have a central role in regulating the coupling of neuronal activity to cerebral glucose metabolism [59,60].

Our results also suggest that glutamatergic dysfunction in the PFC of PCP-treated rats is not limited to the disruption of glutamatergic neurotransmission but also involves the disruption of the metabolic pathways in which glutamate is involved. For example, altered glutamate metabolism may directly contribute to the disruption of the Arginine-Proline metabolic pathway in the PFC of PCP-treated animals. The significant disruption of the Arginine pathway in PCP-treated animals suggests that prolonged

Table 5 Hypergeometric probability of KEGG defined metabolic pathways in bottom cluster in Figure 11

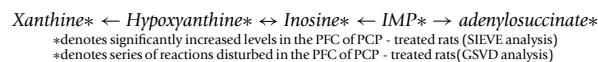
KEGG Path-way Identity	KEGG Pathway	Number of metabolites in cluster(A)	Total number of pathway metabolites detected (B)	Hypergeometric Probability ($P(X) \geq k$)
ko00010	Glycolysis/Gluconeogenesis	1	1	0.184
ko00051	Fructose and Mannose metabolism	1	1	0.184
ko00052	Galactose metabolism	1	1	0.184
ko00071	Fatty acid metabolism	1	1	0.184
ko00230	Purine metabolism	4	13	0.191
ko00260	Glycine, Serine and Threonine metabolism	1	7	0.770
ko00270	Cysteine and Methionine metabolism	1	7	0.770
ko00340	Histidine metabolism	1	5	0.646
ko00410	beta-Alanine metabolism	1	5	0.646
ko00430	Taurine and Hypotaurine metabolism	1	3	0.460
ko00440	Phosphonate and Phosphinate metabolism	1	2	0.335
ko00480	Glutathione metabolism	1	5	0.646
ko00561	Glycerolipid metabolism	1	2	0.335
ko00562	Inositol Phosphate metabolism	1	2	0.335
ko00564	Glycerphospholipid metabolism	2	11	0.642
ko00620	Pyruvate metabolism	1	2	0.335
map00650	Butanoate metabolism	1	4	0.562
ko00730	Thiamine metabolism	1	1	0.184
ko00770	Pantothenate and CoA biosynthesis	1	5	0.646
map00920	Sulphur metabolism	1	3	0.460

Table 5 shows the hypergeometric probability of randomly seeing at least the observed number of metabolites of a given KEGG pathway in the bottom cluster of Figure 11, identified though the GSVD algorithm as being present in control animals but not in PCP-treated animals. There was no evidence for a particular over-abundance of metabolites from any given KEGG pathway in this cluster. Cluster size is 18 metabolites from a total population of 98.

NMDA receptor hypofunction may result in disrupted nitric oxide (NO) signalling in the PFC. There is increasing evidence that NO signalling is directly linked to NMDA receptor activity through regulation of the enzyme nitric oxide synthase (NOS) [61] and that NO signaling contributes to the deficits in cognition that arise from acute NMDA receptor blockade [62,63]. The finding that Citrulline levels, a metabolite in the Arginine-Proline pathway, are significantly decreased in the PFC of the PCP-treated rats in this study further supports the suggestion that NOS activity is altered in the PFC of these animals, as this metabolite is formed by NOS when it releases NO from L-arginine. This suggests that NMDA receptor hypofunction may underlie the decreased NOS activity and protein expression levels reported in the PFC of schizophrenia patients [64] and may contribute to the cognitive deficits seen in this disorder.

In addition to quantitatively defining the specific metabolic pathways altered by experimental manipulation, our

results suggest that the GSVD algorithm can identify discrete series of metabolic reactions altered by experimental manipulation. In this way, while we found no significant evidence to support the widespread disruption of purine metabolism, or the significant disruption of any other KEGG defined metabolic pathway in the bottom cluster as detected using the GSVD, we did find evidence in this cluster to suggest that a specific series of purine reactions were significantly disrupted in the PFC of PCP-treated animals. These disrupted purinergic reactions in the PFC of PCP-treated animals were:



This result suggests that the activity of adenylosuccinate synthase (ADSS), the enzyme responsible for the conversion of IMP to adenylosuccinate, may be significantly increased in the PFC of PCP-treated animals. An increase in the functional activity of this enzyme could

result in both the increased level of adenylosuccinate and the altered balance in the enzyme's downstream metabolites (IMP, Inosine, Hypoxanthine, Xanthine) seen in the PFC of PCP-treated animals. While the influence of prolonged NMDA receptor hypofunction on the functional activity of this specific enzyme remains to be confirmed, and clearly warrants further systematic investigation, the recent finding of altered ADSS gene expression in schizophrenia [65] and the association of ADSS gene polymorphisms with schizophrenia [66] further highlights a potential role for this metabolic pathway in this disorder. In addition, a role for this metabolic pathway in cognition and schizophrenia is supported by the observation that inherited deficiency in the enzyme responsible for the breakdown of adenylosuccinate (ASL) results in mental retardation and autistic features [67,68]. Furthermore, the ASL gene maps to chromosome 22q13.1-q13.2 in humans [69] and these chromosomal loci have been repeatedly linked to schizophrenia [70-72]. The disruption of this metabolic pathway may also contribute to the reduced rate of cerebral glucose metabolism in the PFC of PCP-treated animals [3,4] as ASL deficiency results in hypometabolism in frontal cortical structures [73]. Overall, these results suggest that the potential role of this specific series of metabolic reactions and its enzymes in cognition and schizophrenia warrants further investigation.

Conclusions

This work addresses the scenario where a pair of networks describes two different patterns of connection between a common set of nodes. We argued from first principles that the Generalized Singular Value Decomposition (equation (4)) can form the basis of a very useful computational tool. In practice, we have shown that this new computational network reordering technique was able to identify alterations in metabolic pathways in the PFC of rats treated subchronically with PCP that may contribute to the PFC dysfunction and cognitive deficits seen in these animals. Furthermore, the metabolic pathways identified as being disrupted in the PFC of PCP-treated rats through the application of this new computation technique clearly overlap with those metabolic species known to be disrupted in schizophrenia. Applying this new algorithm in this way also identified novel pathways that may also be relevant to schizophrenia. In this way we identified alterations in glutamate metabolism and metabolic pathways central to glutamatergic neurotransmission, alterations in arginine and proline metabolism and the disruption of a novel series of purine reactions that may contribute to the PFC dysfunction and cognitive deficits seen in schizophrenia.

Methods

Chemicals

The solvents used for the study were purchased from the following sources: Acetonitrile, methanol and chloroform (Fisher Scientific, Leicestershire, UK) and formic acid (VWR, Poole, UK). All chemicals used were of analytical reagent grade. A Direct Q-3[®] water purification system (Millipore, Watford, UK) was used to produce HPLC grade water which was used in all analysis. Standards for 90 common bio-molecules were also purchased which were used to characterize the ZIC-HILIC column (Sigma Aldrich, Dorset UK).

Animals

All experiments were completed using male Lister Hooded rats (Harlan-Olac, UK) housed under standard conditions (21°C, 45-65% humidity, 12-h dark/light cycle (lights on 0600h) with food and drinking water available *ad libitum*). All manipulations were carried out at least 1 week after entry into the facility and all experiments were carried out under the Animals (Scientific Procedures) Act 1986. Animals received either sub-chronic treatment with vehicle (0.9% saline, *i.p.*, $n = 5$) or 2.58mg.kg⁻¹ PCP.HCl (*i.p.*, Sigma Aldrich, UK) once daily for five consecutive days ($n = 5$). At 72 hours after the final drug treatment dose animals were sacrificed and the brain rapidly dissected out and frozen in isopentane (-40°C) and stored at -80°C until sectioning. Frozen brains were sectioned (20 μ M) in the coronal plane in a cryostat (-20°C). Tissue sections from the prefrontal cortex (PFC, Bregma +4.70mm to Bregma +3.20mm) were collected in 4ml glass vials with reference to a stereotactic rat brain atlas [74] and stored at -80°C until further preparation for LC-MS analysis.

Extraction of polar metabolites from brain samples for LC-MS analysis

Extraction of polar metabolites from brain tissue was carried out using the two-step extraction method described previously [75], using methanol, water and chloroform for the optimal extraction of polar metabolites. A hand held homogenizer was used to homogenize the samples once in solution. For preparation of samples for LC-MS analysis 200 μ l of the collected polar extract was added to 600 μ l of 1 : 1 acetonitrile:water solution to produce a final solvent:sample ratio of 3 : 1. The samples were then filtered using Acrodisc 13mm syringe filters with 0.2 μ m nylon membrane (Sigma Aldrich) before LC-MS analysis.

LC-MS analysis of polar metabolites

Experiments were carried out using a Finnigan LTQ Orbitrap (Thermo Fisher, Hemel Hempstead, UK) using

30000 resolution. Analysis was carried out in positive mode over a mass range of 60-1000 m/z . The capillary temperature was set at 250°C and in positive ionization mode the ion spray voltage was 4.5 kV, the capillary voltage 30 V and the tube lens voltage 105 V. The sheath and auxiliary gas flow rates were 45 and 15, respectively (units not specified by manufacturer). A ZIC-HILIC column (5 μm , 150 \times 4.6 mm; HiChrom, Reading, UK) was used in all analysis and a binary gradient method was developed which produced good polar metabolite separation. Solvent A was 0.1% v/v formic acid in HPLC grade water and solvent B was 0.1% v/v formic acid in acetonitrile. A flow rate of 0.3 ml/min. was used and the injection volume was 10 μl . The gradient programme used was 80% B at 0 min. to 50% B at 12 min. to 20% B at 28 min. to 80% B at 37 min., with total run time of 45 minutes. The instrument was externally calibrated before analysis and internally calibrated using lock masses at m/z 83.06037 and m/z 195.08625. Samples were analysed sequentially and the vial tray temperature was set at a constant temperature of 4°C.

Data preparation and analysis

Determination of overt alterations in metabolite levels between experimental groups

The software program Xcalibur (version 2.0) was used to acquire the LC-MS data. The raw Xcalibur data files from version 1.2 (Thermo Fisher, Hemel Hempstead, UK). SIEVE software (Thermo-Fisher Scientific) was used to identify all metabolites affected by drug treatment by calculating a p -value and ratio based on the difference in average intensities of individual peaks, which correspond to different metabolites, between PCP-treated and control animals. A significant difference in the level of each metabolite between groups was set at p -value < 0.05 and/or ratio less than 0.5 for downregulated metabolites and greater than 2 for upregulated metabolites. The ratio is the fold change in average peak intensities from control and treatment groups. For metabolite identification the masses of the polar metabolites were compared to the exact masses of 6000 biomolecules using an in-house developed macro (Excel, Microsoft 2007).

Hypergeometric probability testing

The hypergeometric probability test was used to calculate the probability of finding at least the observed number of metabolites of a given pre-defined metabolic pathway (as defined on the KEGG pathway database) in the clusters identified through the GSVD algorithm, with knowledge of the total number of metabolites present in that pathway detected by LC-MS in these samples. The hypergeometric probability test was used to identify whether any of the KEGG defined metabolic pathways were significantly over-represented in any of

the GSVD identified clusters. In its general form hypergeometric probability allows the calculation of the probability of observing at least (k) metabolites from a given defined KEGG pathway in a defined cluster of metabolites (n) given the total number of metabolites (N) and the total number of metabolites from the pathway in question (m). The probability mass function of hypergeometric distribution is:

$$f(k; N, m, n) = P(X = k) = \frac{\binom{m}{k} \binom{N-m}{n-k}}{\binom{N}{n}}. \quad (9)$$

So here the probability is calculated using the formula

$$P(X \geq k) = \sum_{i=k}^m \frac{\binom{m}{i} \binom{N-m}{n-i}}{\binom{N}{n}}. \quad (10)$$

Significant over-representation of a given functional group in any GSVD defined significant cluster was set by a hypergeometric probability threshold of 0.05.

Additional material

Additional file 1: Table S1 - List of all metabolites detected by LC-MS in the PFC of Control and PCP-treated animals. Table S1 Legend:

The molecular formula and tentative molecular identity for each metabolite detected in the PFC of control and PCP-treated animals is shown. In addition, the KEGG molecular identity and the KEGG metabolic pathways in which a metabolite is involved are also shown. The ratio difference in metabolite concentration and the significance of this change (p -value), as determined by SIEVE analysis (see Methods section), are also shown. Those metabolites found to be significantly different between the two groups are highlighted in bold. The most prominent alterations in KEGG defined metabolic pathways appeared to be in (i) alanine, aspartate and glutamate metabolism (3 metabolites [ko00250]), (ii) phenylalanine, tyrosine and tryptophan metabolism (3 metabolites [ko00360]), (iii) purine metabolism (2 metabolites [ko00230]) and (iv) butanoate metabolism (2 metabolites [ko00650]). KEGG defined metabolic pathways; **ko00250**: Alanine, Aspartate and Glutamate metabolism; **ko00627**: Aminobenzoate degradation; **ko00330**: Arginine and Proline metabolism; **ko00410**: beta-Alanine metabolism; **ko00780**: Biotin metabolism; **map00650**: Butanoate metabolism; **ko04973**: Carbohydrate metabolism; **ko00270**: Cysteine and Methionine metabolism; **ko00071**: Fatty acid metabolism; **ko00051**: Fructose and Manose metabolism; **ko00052**: Galactose metabolism; **ko00471**: Glutamine and Glutamate metabolism; **ko0480**: Glutathione metabolism; **ko00561**: Glycerolipid metabolism; **ko00564**: Glycerophospholipid metabolism; **ko00260**: Glycine, Serine and Threonine metabolism; **ko00010**: Glycolysis/Gluconeogenesis; **ko00340**: Histidine metabolism; **ko00562**: Inositol Phosphate metabolism; **map00300**: Lysine biosynthesis; **ko00310**: Lysine degradation; **ko00430**: Methionine metabolism; **ko04080**: Neuroactive ligand-receptor interaction; **ko00760**: Nicotinate and Nicotinamide metabolism; **ko00190**: Oxidative phosphorylation; **ko00770**: Pantoate and CoA biosynthesis; **ko00550**: Peptidoglycan biosynthesis; **ko00360**: Phenylalanine metabolism; **ko00400**: Phenylalanine, Tyrosine and Tryptophan biosynthesis; **ko00440**: Phosphonate and Phosphinate metabolism; **ko00640**: Propanoate metabolism; **ko00230**: Purine metabolism; **ko00240**: Pyrimidine metabolism; **ko00620**: Pyruvate metabolism; **ko00500**: Starch and Sucrose metabolism; **ko00600**: Sphingolipid metabolism; **ko00920**: Sulphur metabolism; **ko00430**: Taurine and Hypotaurine metabolism; **ko00730**: Thiamine metabolism; **ko00380**: Tryptophan metabolism; **ko00350**: Tyrosine metabolism; **ko00400**: Tyrosine and Tryptophan

biosynthesis; **ko00290**: Valine, Leucine and Isoleucine biosynthesis; **ko00280**: Valine, Leucine and Isoleucine degradation. NA denotes a metabolite not associated with a KEGG compound ID or KEGG pathway.

Additional file 2: 98 × 98 matrix of between metabolite correlations in the PFC of control animals. The 98 × 98 matrix of the Pearson's correlation coefficients (Fisher z-transformed) between all metabolites detected in the prefrontal cortex of control (saline-treated) animals by LC-MS analysis is shown.

Additional file 3: 98 × 98 matrix of between metabolite correlations in the PFC of PCP-treated animals. The 98 × 98 matrix of the Pearson's correlation coefficient (Fisher z-transformed) between all metabolites detected in the prefrontal cortex of PCP-treated animals by LC-MS analysis is shown.

Additional file 4: Table S2 - Table showing the axes labels in Figures 9, 10 and 11. In Table S2 the position of each metabolite in the original ordering (Figure 4) is shown. In the columns for Figures 5 and 6, the corresponding numbers indicating the new position of each metabolite (node) in the matrix when reordered by the first column of X^{-T} and the final column of X^{-T} , respectively, is shown.

Acknowledgements

XX is supported by Engineering and Physical Sciences Research Council grant EP/E049370/1.

ND is supported by Psychiatric Research Institute in Glasgow (PsyRING), a joint initiative between the Universities of Glasgow and Strathclyde and the National Health Service of Greater Glasgow and Clyde. LM and DGW are supported by the Scottish Universities Life Sciences Alliance (SULSA). DJH is supported by Engineering and Physical Sciences Research Council grant EP/E049370/1 and Medical Research Council grant G0601353.

Author details

¹Department of Mathematics and Statistics, University of Strathclyde, Glasgow, G1 1XH, Scotland, UK. ²Psychiatric Research Institute of Neuroscience in Glasgow (PsyRING), Universities of Glasgow and Strathclyde, G12 8QQ, UK. ³Strathclyde Institute of Pharmacy and Biomedical Sciences (SIPBS), University of Strathclyde, Glasgow, G4 0NR, UK. ⁴Center for Neuroscience, University of Strathclyde (CeNsUS), Glasgow, G4 0NR, UK. ⁵Institute of Neuroscience and Psychology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, G12 8QQ, UK.

Authors' contributions

All authors contributed extensively to the work presented in this paper. DJH and XX conceived and analyzed the computational algorithm, designed and performed the synthetic tests and wrote the description of this material. XX applied the algorithm to metabolic networks. ND performed the *in vivo* experiments, analyzed the metabolomic data and wrote the description of this material. BJM and JAP conceived the subchronic PCP model. LM and DGW conducted the metabolomics experiments. All authors discussed the results, interpreted the data and have read and approved the final version of the manuscript.

Received: 29 October 2010 Accepted: 16 May 2011

Published: 16 May 2011

References

- Davidson LL, Heinrichs RW: Quantification of frontal and temporal lobe brain-imaging findings in schizophrenia: a meta-analysis. *Psychiatry Research: Neuroimaging* 2003, **122**:69-87.
- Hill K, Mann L, Laws KR, Stephenson CME, Nimmo-Smith I, McKenna PJ: Hypofrontality in schizophrenia: a meta-analysis of functional imaging studies. *Acta Psychiatrica Scandinavica* 2004, **110**:243-256.
- Cochran SM, Kennedy M, McKerchar CE, Steward LJ, Pratt JA, Morris BJ: Induction of Metabolic Hypofunction and Neurochemical Deficits after Chronic Intermittent Exposure to Phencyclidine: Differential Modulation by Antipsychotic Drugs. *Neuropsychopharmacology* 2003, **28**:265-275.

- Dawson N, Thompson RJ, McVie A, Thomson DM, Morris BJ, Pratt JA: Modafinil reverses phencyclidine (PCP)-induced deficits in cognitive flexibility, cerebral metabolism and functional brain connectivity. *Schizophrenia Bulletin* .
- Egerton A, Reid L, McGregor S, Cochran SM, Morris BJ, Pratt JA: Subchronic and chronic PCP treatment produces temporally distinct deficits in attentional set shifting and prepulse inhibition in rats. *Psychopharmacology* 2008, **198**:37-49.
- Egerton A, Reid L, McKerchar CE, Morris BJ, Pratt JA: Impairment in perceptual attentional set-shifting following PCP administration: a rodent model of set-shifting deficits in schizophrenia. *Psychopharmacology* 2005, **179**:77-84.
- Steward LJ, Kennedy MD, Morris BJ, Pratt JA: The atypical antipsychotic drug clozapine enhances chronic PCP-induced regulation of prefrontal cortex 5-HT_{2A} receptors. *Neuropharmacology* 2004, **47**:527-537.
- Fiehn O: Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics* 2001, **2**:155-168.
- Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB: Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in Biotechnology* 2004, **22**:245-252.
- Goodacre R, York EV, Heald JK, Scott IM: Chemometric discrimination of unfractionated plant extracts analyzed by electrospray mass spectrometry. *Phytochemistry* 2003, **62**:859-863.
- Oliver SG, Winson MK, Kell DB, Baganz F: Systematic functional analysis of the yeast genome. *Trends in Biotechnology* 1998, **16**:373-378.
- Kaddurah-Daouk R, Krishnan KRR: Metabolomics: A Global Biochemical Approach to the Study of Central Nervous System Diseases. *Neuropsychopharmacology* 2009, **34**:173-186.
- Quinones MP, Kaddurah-Daouk R: Metabolomics tools for identifying biomarkers for neuropsychiatric diseases. *Neurobiology of Disease* 2009, **35**:165-176.
- Greenberg N, Grassano A, Thambisetty M, Lovestone S, Legido-Quigley C: A proposed metabolic strategy for monitoring disease progression in Alzheimer's disease. *Electrophoresis* 2009, **30**:1235-1239.
- Chen C, Gonzalez FJ, Idle JR: LC-MS-based metabolomics in drug metabolism. *Drug Metab Rev* 2007, **39**:581-597.
- Lindon JC, Holmes E, Nicholson JK: Metabolomics in pharmaceutical research and development. *FEBS Journal* 2007, **274**:1140-1151.
- Robertson DG: Metabolomics in Toxicology: A Review. *Toxicological Sciences* 2005, **85**:809-822.
- Cheng KK, Benson GM, Grimsditch DC, Reid DG, Connor SC, Griffin JL: Metabolomic study of the LDL receptor null mouse fed a high-fat diet reveals profound perturbations in choline metabolism that are shared with ApoE null mice. *Physiological Genomics* 2010, **41**:224-231.
- Dettmer K, Aronov PA, Hammock BD: Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews* 2007, **26**:51-78.
- Lu W, Bennett BD, Rabinowitz JD: Analytical strategies for LC-MS-based targeted metabolomics. *Journal of Chromatography B* 2008, **871**:236-242.
- Wilson ID, Plumb R, Granger J, Major H, Williams R, Lenz EM: HPLC-MS-based methods for the study of metabolomics. *Journal of Chromatography B* 2005, **817**:67-76.
- Ohta D, Kanaya S, Suzuki H: Application of Fourier-transform ion cyclotron resonance mass spectrometry to metabolic profiling and metabolite identification. *Current Opinion in Biotechnology* 2010, **21**:35-44.
- Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R: The Orbitrap: a new mass spectrometer. *Journal of Mass Spectrometry* 2005, **40**:430-443.
- Calvano CD, Zamboni CG, Jensen ON: Assessment of lectin and HILIC based enrichment protocols for characterization of serum glycoproteins by mass spectrometry. *Journal of Proteomics* 2008, **71**:304-317.
- Cubbon S, Bradbury T, Wilson J, Thomas-Oates J: Hydrophilic Interaction Chromatography for Mass Spectrometric Metabolomic Studies of Urine. *Analytical Chemistry* 2007, **79**:8911-8918.
- Gika HG, Theodoridis GA, Wilson ID: Hydrophilic interaction and reversed-phase ultra-performance liquid chromatography TOF-MS for metabolomic analysis of Zucker rat urine. *Journal of Separation Science* 2008, **31**(9):1598-1608.
- Idborg H, Zamani L, Edlund PO, Schuppe-Koistinen I, Jacobsson SP: Metabolic fingerprinting of rat urine by LC/MS: Part 1. Analysis by

- hydrophilic interaction liquid chromatography-electrospray ionization mass spectrometry. *Journal of Chromatography B* 2005, **828**:9-13.
28. Kind T, Tolstikov V, Fiehn O, Weiss RH: **A comprehensive urinary metabolomic approach for identifying kidney cancer.** *Analytical Biochemistry* 2007, **363**:185-195.
 29. Nordstrom A, Want E, Northen T, Lehtio J, Siuzdak G: **Multiple Ionization Mass Spectrometry Strategy Used To Reveal the Complexity of Metabolomics.** *Analytical Chemistry* 2008, **80**:421-429.
 30. Paek IB, Moon Y, Ji HY, Kim HH, Lee HW, Lee YB, Lee HS: **Hydrophilic interaction liquid chromatography-tandem mass spectrometry for the determination of levosulpiride in human plasma.** *Journal of Chromatography B* 2004, **809**:345-350.
 31. Zhang X, Rauch A, Lee H, Xiao H, Rainer G, Logothetis NK: **Capillary hydrophilic interaction chromatography/mass spectrometry for simultaneous determination of multiple neurotransmitters in primate cerebral cortex.** *Rapid Communications in Mass Spectrometry* 2007, **21**:3621-3628.
 32. Dunn WB, Bailey NJC, Johnson HE: **Measuring the metabolome: current analytical technologies.** *The Analyst* 2005, **130**:606-625.
 33. Spagou K, Tsoukali H, Raikos N, Gika H, Wilson ID, Theodoridis G: **Hydrophilic interaction chromatography coupled to MS for metabolomic/metabolomic studies.** *Journal of Separation Science* 2010, **33**:716-727.
 34. Albert R, Barabási AL: **Statistical mechanics of complex networks.** *Reviews of Modern Physics* 2002, **74**:47-97.
 35. Newman MEJ: **The structure and function of complex networks.** *SIAM Review* 2003, **45**:167-256.
 36. Strogatz SH: **Exploring complex networks.** *Nature* 2001, **410**:268-276.
 37. Shi J, Malik J: **Normalized Cuts and Image Segmentation.** *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2000, **22**:888-905.
 38. Higham DJ: **Spectral reordering of a range-dependent weighted random graph.** *IMA J. Numer. Anal* 2005, **25**:443-457.
 39. Estrada E: **Protein bipartivity and essentiality in the yeast protein-protein interaction network.** *J. Proteome Res* 2006, **5**:2177-2184.
 40. Estrada E, Higham DJ, Hatano N: **Communicability and multipartite structures in complex networks at negative absolute temperatures.** *Phys. Rev. E* 2008, **77**:026102.
 41. Morrison JL, Breitling R, Higham DJ, Gilbert DR: **A Lock-and-Key Model for Protein-Protein Interactions.** *Bioinformatics* 2006, **22**:2012-2019.
 42. Thomas A, Cannings R, Monk NAM, Cannings C: **On the structure of protein-protein interaction networks.** *Biochemical Soc. Trans* 2003, **31**:1491-1496.
 43. Alter O, Brown PO, Botstein D: **Generalized Singular Value Decomposition for Comparative Analysis of Genome-Scale Expression Datasets of Two Different Organisms.** *Proceedings of the National Academy of Sciences* 2003, **100**:3351-3356.
 44. Schreiber A, Shirley NJ, Burton RA, Fincher GB: **Combining transcriptional datasets using the generalized singular value decomposition.** *BMC Bioinformatics* 2008, **9**:335.
 45. Higham DJ, Kalna G, Kibble M: **Spectral clustering and its use in bioinformatics.** *J. Computational and Applied Math* 2007, **204**:25-37.
 46. Fiedler M: **A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory.** *Czechoslovak Mathematical Journal* 1975, **25**:619-633.
 47. Atkins JE, Boman EG, Hendrickson B: **A spectral algorithm for seriation and the consecutive ones problem.** *SIAM Journal on Computing* 1998, **28**:297-310.
 48. Robles-Kelly A, Hancock ER: **Graph Edit Distance from Spectral Seriation.** *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2005, **27**:365-378.
 49. Golub GH, Van Loan CF: *Matrix Computations* Baltimore: Johns Hopkins University Press; 1996.
 50. Strang G: *Computational Science and Engineering* Wellesley-Cambridge Press; 2008.
 51. Estrada E: **Topological structural classes of complex networks.** *Physical Review E* 2007, **75**:016103.
 52. Xiao X: **Complex Networks and the Generalized Singular Value Decomposition.** *PhD thesis* University of Strathclyde, Department of Mathematics and Statistics; 2010.
 53. Murai R, Noda Y, Matsui K, Kamei H, Mouri A, Matsuba K, Nitta A, Furukawa H, Nabeshima T: **Hypofunctional glutamatergic neurotransmission in the prefrontal cortex is involved in the emotional deficit induced by repeated treatment with phencyclidine in mice: Implications for abnormalities of glutamate release and NMDA-CaMKII signaling.** *Behavioural Brain Research* 2007, **180**:152-160.
 54. Hahn CG, Wang HY, Cho DS, Talbot K, Gur RE, Berrettini WH, Bakshi K, Kamins J, Borgmann-Winter KE, Siegel SJ, Gallop RJ, Arnold SE: **Altered neuregulin 1-erbB4 signaling contributes to NMDA receptor hypofunction in schizophrenia.** *Nature Medicine* 2006, **12**:824-828.
 55. Lewis DA, Moghaddam B: **Cognitive Dysfunction in Schizophrenia: Convergence of gamma-Aminobutyric Acid and Glutamate Alterations.** *Arch Neurol* 2006, **63**:1372-1376.
 56. Javitt DC, Balla A, Burch S, Suckow R, Xie S, Sershen H: **Reversal of Phencyclidine-Induced Dopaminergic Dysregulation by N-Methyl-D-Aspartate Receptor/Glycine-site Agonists.** *Neuropsychopharmacology* 2003, **29**:300-307.
 57. Sershen H, Balla A, Aspromonte JM, Xie S, Cooper TB, Javitt DC: **Characterization of interactions between phencyclidine and amphetamine in rodent prefrontal cortex and striatum: Implications in NMDA/glycine-site-mediated dopaminergic dysregulation and dopamine transporter function.** *Neurochemistry International* 2008, **52**:119-129.
 58. Yang CR, Svensson KA: **Allosteric modulation of NMDA receptor via elevation of brain glycine and d-serine: The therapeutic potentials for schizophrenia.** *Pharmacology & Therapeutics* 2008, **120**:317-332.
 59. Magistretti PJ: **Role of glutamate in neuron-glia metabolic coupling.** *Am J Clin Nutr* 2009, **90**:875S-880S.
 60. Patel AB, de Graaf RA, Mason GF, Rothman DL, Shulman RG, Behar KL: **The contribution of GABA to glutamate/glutamine cycling and energy metabolism in the rat cortex in vivo.** *Proceedings of the National Academy of Sciences* 2005, **102**:5588-5593.
 61. Rameau GA, Tukey DS, Garcin-Hosfield ED, Titcombe RF, Misra C, Khatri L, Getzoff ED, Ziff EB: **Biphasic Coupling of Neuronal Nitric Oxide Synthase Phosphorylation to the NMDA Receptor Regulates AMPA Receptor Trafficking and Neuronal Cell Death.** *The Journal of Neuroscience* 2007, **27**:3445-3455.
 62. Fejgin K, Pålsson E, Wass C, Svensson L, Klamer D: **Nitric Oxide Signaling in the Medial Prefrontal Cortex is Involved in the Biochemical and Behavioral Effects of Phencyclidine.** *Neuropsychopharmacology* 2008, **33**:1874-1833.
 63. Wass C, Svensson L, Fejgin K, Pålsson E, Archer T, Engel JA, Klamer D: **Nitric oxide synthase inhibition attenuates phencyclidine-induced disruption of cognitive flexibility.** *Pharmacology Biochemistry and Behavior* 2008, **89**:352-359.
 64. Xing G, Chavko M, Zhang LX, Yang S, Post RM: **Decreased calcium-dependent constitutive nitric oxide synthase (cNOS) activity in prefrontal cortex in schizophrenia and depression.** *Schizophrenia Research* 2002, **58**:21-30.
 65. Tsuang MT, Nossova N, Yager T, Tsuang MM, Guo SC, Shyu KG, Glatt SJ, Liew C: **Assessing the validity of blood-based gene expression profiles for the classification of schizophrenia and bipolar disorder: A preliminary report.** *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 2005, **133B**:1-5.
 66. Zhang F, Xu Y, Liu P, Fan H, Huang X, Sun G, Song Y, Sham P: **Association analyses of the interaction between the ADSS and ATM genes with schizophrenia in a Chinese population.** *BMC Medical Genetics* 2008, **9**:119-125.
 67. Spiegel EK, Colman RF, Patterson D: **Adenylosuccinate lyase deficiency.** *Molecular Genetics and Metabolism* 2006, **89**(1-2):19-31.
 68. Stone RL, Aimi J, Barshop BA, Jaeken J, Van den Berghe G, Zalkin H, Dixon JE: **A mutation in adenylosuccinate lyase associated with mental retardation and autistic features.** *Nature Genetics* 1992, **1**:59-63.
 69. Fon E, Demczuk S, Delattre O, Thomas G, Rouleau G: **Mapping of the human adenylosuccinate lyase (ADSL) gene to chromosome 22q13.1-13.2.** *Cytogenetics and Cell Genetics* 1993, **64**:201-203.
 70. Jorgensen T, Børglum A, Mors O, Wang A, Pinaud M, Flint T, Dahl H, Vang M, Kruse T, Ewald H: **Search for common haplotypes on chromosome 22q in patients with schizophrenia or bipolar disorder from the Faroe Islands.** *American Journal of Medical Genetics* 2002, **114**:245-252.
 71. Pulver AE, Karayiorgou M, Wolyniec PS, Lasseter VK, Kasch L, Nestadt G, Antonarakis S, Housman D, Kazazian HH, Meyers D, Ott J, Lamacz M, Liang KY, Hanfelt J, Ullrich G, DeMarchi N, Ramu E, McHugh PR, Adler L,

- Thomas M, Carpenter WT, Manschreck T, Gordon CT, Kimberland M, Babb R, Puck J, Childs B: **Sequential strategy to identify a susceptibility gene for schizophrenia: Report of potential linkage on chromosome 22q12-q13.1: Part 1.** *American Journal of Medical Genetics* 1994, **54**:36-43.
72. Saleem Q, Dash D, Gandhi C, Kishore A, Benegal V, Sherrin T, Mukherjee O, Jain S, Brahmachari SK: **Association of CAG repeat loci on chromosome 22 with schizophrenia and bipolar disorder.** *Molecular Psychiatry* 2001, **6**:694-700.
73. De Volder AG, Jaeken J, Den Berghe GV, Bol A, Michel C, Cogneau M, Goffinet AM: **Regional Brain Glucose Utilization in Adenylosuccinase-Deficient Patients Measured by Positron Emission Tomography.** *Pediatric Research* 1998, **24**:238-242.
74. Paxinos G, Watson C: *The Rat Brain in Stereotaxic Coordinates* Academic Press; 1998.
75. Wu H, Southam AD, Hines A, Viant MR: **High-throughput tissue extraction protocol for NMR- and MS-based metabolomics.** *Analytical Biochemistry* 2008, **372**:204-212.

doi:10.1186/1752-0509-5-72

Cite this article as: Xiao *et al.*: Exploring metabolic pathway disruption in the subchronic phencyclidine model of schizophrenia with the Generalized Singular Value Decomposition. *BMC Systems Biology* 2011 **5**:72.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

