



OPEN

## A system biology approach based on metabolic biomarkers and protein–protein interactions for identifying pathways underlying schizophrenia and bipolar disorder

Md. Altaf-Ul-Amin<sup>1</sup>✉, Kazuhisa Hirose<sup>1</sup>, João V. Nani<sup>2,3</sup>, Lucas C. Porta<sup>2</sup>, Ljubica Tasic<sup>4</sup>, Shaikh Farhad Hossain<sup>1</sup>, Ming Huang<sup>1</sup>, Naoaki Ono<sup>1</sup>, Mirian A. F. Hayashi<sup>2,3</sup>✉ & Shigehiko Kanaya<sup>1</sup>

Mental disorders (MDs), including schizophrenia (SCZ) and bipolar disorder (BD), have attracted special attention from scientists due to their high prevalence and significantly debilitating clinical features. The diagnosis of MDs is still essentially based on clinical interviews, and intensive efforts to introduce biochemical based diagnostic methods have faced several difficulties for implementation in clinics, due to the complexity and still limited knowledge in MDs. In this context, aiming for improving the knowledge in etiology and pathophysiology, many authors have reported several alterations in metabolites in MDs and other brain diseases. After potentially fishing all metabolite biomarkers reported up to now for SCZ and BD, we investigated here the proteins related to these metabolites in order to construct a protein–protein interaction (PPI) network associated with these diseases. We determined the statistically significant clusters in this PPI network and, based on these clusters, we identified 28 significant pathways for SCZ and BDs that essentially compose three groups representing three major systems, namely stress response, energy and neuron systems. By characterizing new pathways with potential to innovate the diagnosis and treatment of psychiatric diseases, the present data may also contribute to the proposal of new intervention for the treatment of still unmet aspects in MDs.

Schizophrenia (SCZ) and bipolar disorder (BD) are severe debilitating mental disorders (MDs), as both are associated with cognitive impairments, and altered behavior, mood and perceptions. Together, these MDs affect around 100 million people worldwide, irrespective of nationality, ethnic origin, or socioeconomic status<sup>1</sup>. SCZ is characterized by a set of positive and negative symptoms and, according to the 5th version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V), at least two or more of these symptoms need to be present in an individual for the diagnosis. In turn, BD is a chronic mood disorder often characterized by the fluctuations between mania and depressive episodes, and due to the complex mood alterations, the clinical misdiagnosis in BD is common<sup>2</sup>. In addition, other MDs, such as major depressive disorders, also shares several common symptoms and specific endophenotypes with SCZ and BD<sup>3–5</sup>. In fact, pathophysiology of all these MDs are mainly centered in the hypothesis of alterations in dopamine homeostasis and signaling<sup>6,7</sup>, while gamma-aminobutyric acid (GABA) is the principal inhibitory neurotransmitter found altered in SCZ<sup>8–11</sup>. SCZ and BD share several symptoms and, up to a certain degree, also share similar pharmacological interventions. For instance,

<sup>1</sup>Nara Institute of Science and Technology, Ikoma, Nara 630-0192, Japan. <sup>2</sup>Department of Pharmacology, Escola Paulista de Medicina (EPM), Universidade Federal de São Paulo (UNIFESP), São Paulo, SP, Brazil. <sup>3</sup>National Institute for Translational Medicine (INCT-TM, CNPq/FAPESP/CAPEs), Ribeirão Preto, Brazil. <sup>4</sup>Chemical Biology Laboratory, Department of Organic Chemistry, Institute of Chemistry, Universidade Estadual de Campinas (Unicamp), Campinas, SP, Brazil. ✉email: amin-m@is.naist.jp; mhayashi@unifesp.br

the second-generation antipsychotics (SGAs) were primarily developed for the treatment of the positive and negative symptoms of SCZ<sup>12</sup>, while they are also currently used as an alternative to lithium for the suppression of the main symptoms in BD<sup>13</sup>.

SCZ and BD are highly polygenic diseases, with many associated genetic variants with small effects, as demonstrated by several genome-wide association studies (GWAS)<sup>14–17</sup>. Moreover, several genetic-based studies have shown a limited contribution to support the diagnosis and/or the characterization of pathways underlying these major MDs, due to the well-recognized pleiotropic features and small size effect of each gene<sup>18,19</sup>. Therefore, none of genetic studies was adequate to reveal the mechanisms underlying the etiology and/or pathophysiology of SCZ and/or BD.

In this context, proteomics and metabolomics are now recognized as valuable tools for the characterization of any disease. Although proteins are the functional actors that are responsible for the generation or catabolism of all metabolites, identification of metabolic processes characteristically active or inactive in any specific healthy and/or pathological conditions may uncover pathways important for the understanding of disease mechanisms. Interestingly, the separation of SCZ and healthy control (HC) individuals, by metabolomic analysis was demonstrated to be possible by employing different methods, such as proton NMR (<sup>1</sup>H-NMR) or mass spectrometry (MS)<sup>15,20–22</sup>, which also highlighted that the metabolic changes can be detected in different biological samples, including urine, blood or cerebrospinal fluid (CSF)<sup>23,24</sup>. The separation of BD subjects treated with SGAs, BD subjects treated with lithium, and SCZ subjects treated with SGAs was possible by employing serum metabolomic studies, suggesting also the specificity of the pharmacometabolome to assist the diagnosis in clinics<sup>25,26</sup>. In addition, a set of metabolites that allowed the separation of BD and SCZ patients from HCs, and that could discriminate SCZ from BD and crack users, were recently identified by us and others<sup>15,20,25–28</sup>. These studies may represent a good example of how the metabolomic studies can bring new information about pathways and metabolic alterations, which might not be detectable by genetic studies. However, in these studies, the pathways identified for each individual metabolite could not be clearly associated with the etiology or pathophysiology of SCZ and BD.

In the current study, we applied a systems level approach by exploring together the altered metabolites in SCZ and BD, aiming to have new insights into the underlying pathophysiology of these MDs. A systems level approach is likely to compensate some missing information and discard noisy data from the process. For this purpose, we identified the genes/proteins that are closely related with the biomarker metabolites and which are connected through reliable protein–protein interactions (PPIs). After constructing a PPI network related to SCZ and BD, we applied a graph-clustering algorithm DPCLUSO to determine the clusters in the network. Then, we utilized the statistically significant clusters to identify important and common pathways underlying SCZ and BD.

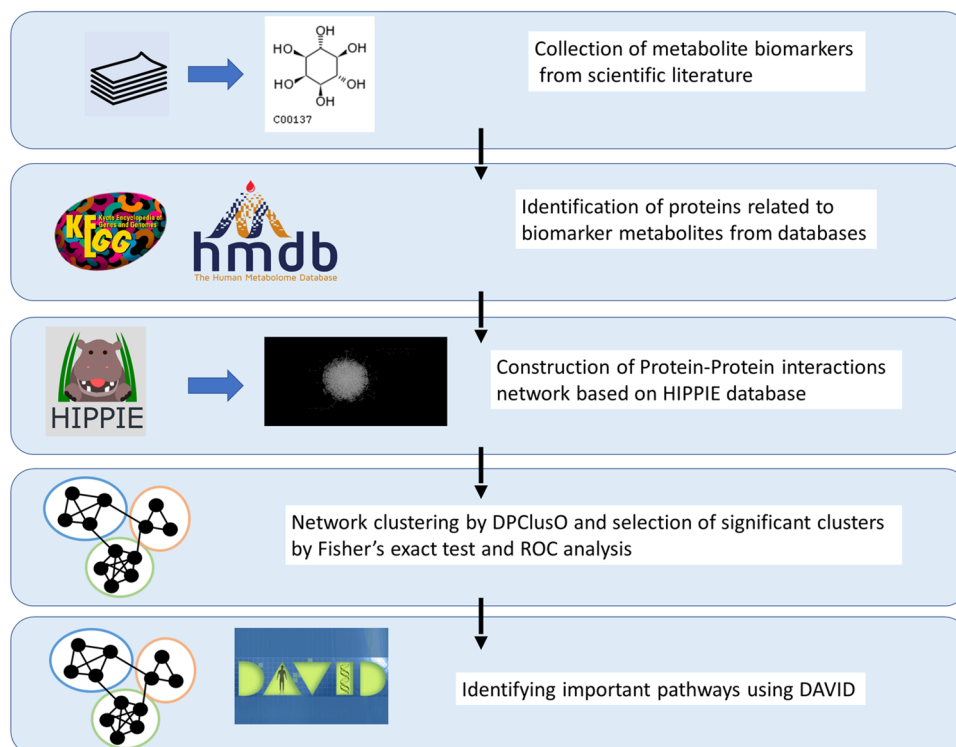
## Methods and results

The approach adopted in the present work is illustrated in the flowchart (Fig. 1), consisting of five major steps. The methods and corresponding results obtained in each step of the flowchart are discussed below in five separate sections, according to the flow order.

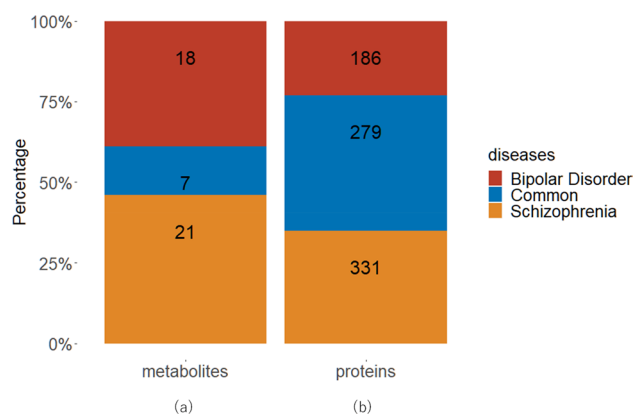
**Data collection of biomarkers.** Biomarker metabolites were collected by searching the most relevant papers related to the target diseases, namely SCZ and/or BD. For this present study, we collected a total of 46 biomarkers, from which 28 were related to SCZ, 25 were related to BD and 7 were common to these both MDs. The proportions of SCZ and BD unique and common metabolites are shown in Fig. 2a.

These biomarkers were mostly detected by NMR, MS, or ion cyclotron type of experiments by using human biological samples, such as serum, blood etc. The collection of these two sets consisting of 46 biomarker metabolites includes: acetate, N-acetyl-D-mannosamine, 2,3-diphospho-D-glyceric-acid, alpha-ketoglutaric-acid, N-acetyl-L-alanine, arginine, choline, formate, glutamate, amygdalin, isocitric acid, myo-inositol, N-acetyl-glutamic-acid, phenylalanine, propionate, pyruvate, serine, beta-alanine, N-acetyl-L-phenyl-alanine, lipoamide, alpha-ketoisovaleric acid (MOA)\*, L-glutamine, acetate, N-acetyl aspartyl-glutamic acid (NAAG), lactate, phosphocholine, alanine, citrate, cystine, eicosanoic acid, glucose, glycerate, β-hydroxybutyrate, pyroglutamic acid, sorbitol, taurine, tocopherol-alpha, uridine, L-threonine, adenine, glycine, adenosine, GABA, mannitol, pantothenate, 3-methyl-2-oxobutanoic acid, guanine. We added these data to KNApSack biomarker database in which the biomarkers of many other diseases have also been collected<sup>29–31</sup>.

**Collection of related proteins.** We adopted two methods to collect SCZ and BD proteins related to the biomarkers reported for these MDs. Firstly, we employed the R package hmdbQuery<sup>32</sup>, which can provide pairwise associations between the metabolites and genes/proteins based on Human Metabolome Database (HMDB)<sup>33</sup>. Secondly, we utilized the metabolic pathway maps of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database<sup>34</sup>. The enzymes up to the path length of two, from each target biomarker, were selected. Here, we illustrate the process using an example. Figure 3 shows a part of the KEGG pathway (human) concerning L-glutamate, which is the synonym for the biomarker Glutamate with KEGG ID C00025. In Fig. 3, we can see 4 branches connected to L-glutamate. All enzymes up to the path length of two, which include GLUD1, GLUD2, NIT2, ALDH4A1, ABAT, GFPT1, GFPT2, GLUL, GLS, GLS2, PPAT and GAD1, were all considered in our study. However, as for instance, CPS wasn't selected, because it was considered out of the proposed protocol, as it was beyond the path length of two. We repeated this process for the 46 different biomarkers considered for this study (as listed above), and then we finally constructed the collection of these 46 sets of enzymes/proteins. Thus, we identified 331 + 279 = 610 and 186 + 279 = 495 proteins related to SCZ and BD, respectively, from which 279 proteins were common to both diseases (as shown in Fig. 2b), also possibly implying that both diseases are closely related. Therefore, for the sake of the system level analysis, it was reasonable to consider these two dis-



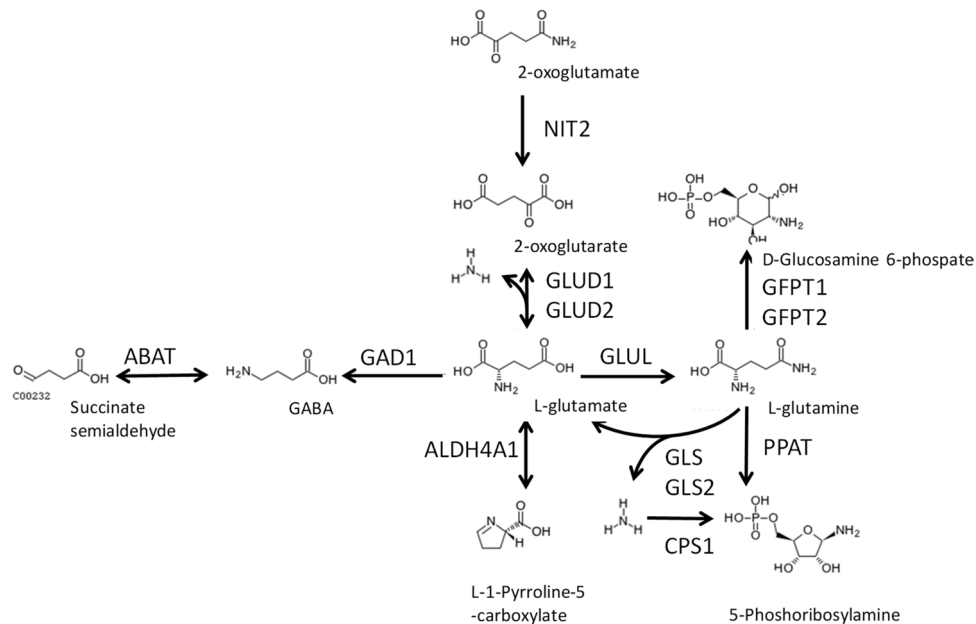
**Figure 1.** Flowchart showing each steps of the present research. The flowchart shows the analysis performed in the present work, starting from fishing out the most relevant metabolite biomarkers from online databases, such as MEDLINE/PubMed (maintained by The United States National Library of Medicine at the National Institutes of Health). Then, the analysis was performed using Kyoto encyclopedia of genes and genomes (KEGG), human metabolome database (HMDB), Human integrated protein–protein interaction rEference (HIPPIE), DAVID (The database for annotation, visualization and integrated discovery) and a network clustering algorithm (DPCLUSO).



**Figure 2.** Proportion of unique and common (a) metabolites, (b) proteins regarding schizophrenia (SCZ) and bipolar disorder (BD). The metabolites and proteins present in BD, SCZ and common to both diseases are indicated by red, orange and blue colors respectively.

eases together for the pathway analysis. Such integrated analysis is likely to allow the identification of common and other novel significant pathways for these two diseases.

**Construction of PPI network relevant to disease.** SCZ and BD related PPI network was constructed using the human integrated protein–protein interaction rEference (HIPPIE) database, in which each interaction is characterized by a score value. The collection of proteins sets identified and depicted in Fig. 2b was considered as  $S$ . We selected all interactions regardless of the score value between proteins  $a$  and  $b$ , as such that  $a \in S$  and  $b \in$



**Figure 3.** Partial L-glutamate pathway. *GLUD1* glutamate dehydrogenase 1, *GLUD2* glutamate dehydrogenase 2, *NIT2* omega-amidase NIT2, *ALDH4A1* aldehyde dehydrogenase 4 family member A1, *ABAT* 4-aminobutyrate aminotransferase, *GFPT1* glutamine-fructose-6-phosphate transaminase 1, *GFPT2* glutamine-fructose-6-phosphate transaminase 2, *GLUL* glutamine synthetase, *GLS* glutaminase, *GLS2* glutaminase 2, *PPAT* phosphoribosyl pyrophosphate amidotransferase, *GAD1* glutamate decarboxylase 1, *CPS1* carbamoyl-phosphate synthase 1, *GABA* gamma-aminobutyric acid.

S. In addition, we selected the interactions with score value  $> 0.7$  between proteins  $a$  and  $b$ , as such that  $a \in S$  and  $b \notin S$ . Then, we identified 3,233 interactions for BD, 4266 interactions for SCZ, and 1904 common interactions for SCZ and BD. In total, these 5595 interactions among a total of 3184 proteins composed our disease related PPI network. Additionally, in order to inspect the global network properties of this network, we employed Cytoscape tool, which is an open-source tool for visualizing and analyzing networks<sup>35</sup>. The degree distribution of this network is of power-law type. Other global topological properties of this network are as follow: clustering coefficient is 0.066, characteristic path length is 4.973, and diameter is 12. Networks with power law degree distribution are so called scale-free networks<sup>36</sup>. Scale-free networks are also small-world networks, if their average path length is small, and if diameter increases logarithmically with the number of vertices<sup>37</sup>. Therefore, our network consisting of 3184 nodes with a power law degree distribution, average path length 4.973 and diameter 12, can be considered as a small world network. These properties are consistent with a PPI network in general.

**Network clustering and selection of significant clusters.** After constructing the disease relevant PPI network, clustering was performed using the DPclusO algorithm<sup>38–41</sup>. The DPclusO algorithm generates overlapping clusters and ensures coverage, i.e. each node goes to at least one cluster. We hypothesize that clustering of a disease relevant PPI network may help the isolation of systems with disease-related properties. Therefore, statistically significant PPI clusters that are enriched for SCZ and BD related proteins could be used to predict novel genes and pathways. Clustering was performed 9 times using input densities as 0.1, 0.2, ..., up to 0.9. Table 1 shows the statistics data, i.e., the number of clusters, size of the biggest cluster, average cluster size and the number of significant clusters corresponding to 9 sets of the generated clusters. As expected, smaller density values resulted in larger and fewer number of clusters. The enrichment of proteins related to SCZ and BD was assessed in each identified cluster by Fisher's exact test  $p$ -values. Additionally,  $p$ -values were corrected by Bonferroni and Hochberg False Discovery Rate (FDR).

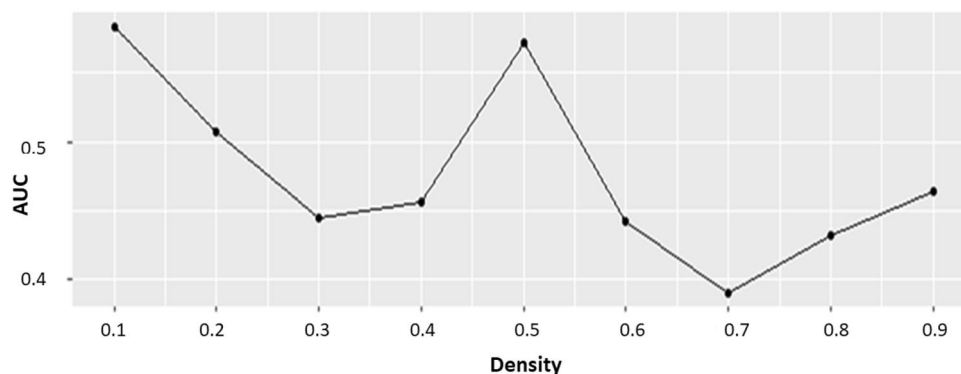
With 9 different input densities, DPclusO generated 9 sets of clusters. In other words, the disease relevant PPI network was divided into clusters in 9 different ways. To assess which set of clusters was more useful for our purpose, we performed a receiver-operating characteristic (ROC) analysis. For ROC analysis, we assigned the SScore (Significance Score)<sup>42</sup>, to each gene, based on the  $p$ -values of the clusters to which they belong to. SScore of a protein is defined as following:

$$\text{SScore} = -\log(\text{FDR}).$$

We assigned a SScore value to each protein. However, as DPclusO performs overlapping clustering, each protein may have more than one SScores, from which we considered only the highest SScore for the ROC analysis. SCZ and BD related genes/proteins were downloaded from the DisGeNet database, and we considered the set collection of DisGeNet data and the extracted set of SCZ and BD related proteins as the true positive proteins

Density	Total clusters	Max size	Average size	Significant clusters
0.1	315	72	17.59683	75
0.2	531	38	10.00188	56
0.3	816	25	6.530637	33
0.4	1053	17	5.361823	29
0.5	1382	13	4.293054	25
0.6	1841	11	3.260185	2
0.7	2724	8	2.104993	49
0.8	2744	7	2.097303	43
0.9	2764	7	2.068017	61

**Table 1.** Results of each DPplusO clustering.



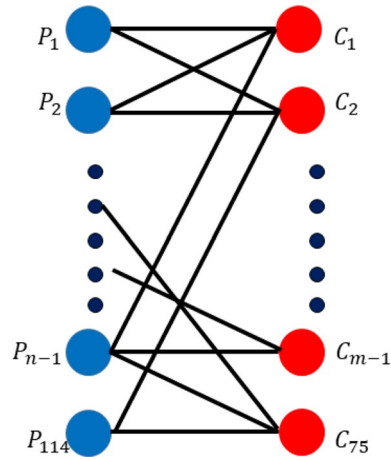
**Figure 4.** The area under the curve (AUC) values for 9 different input densities used for clustering by DPplusO.

related to these MDs. The ROC curve was created by selecting a series of threshold SScore values to generate the true positive rates (TPR) and false positive rates (FPR).

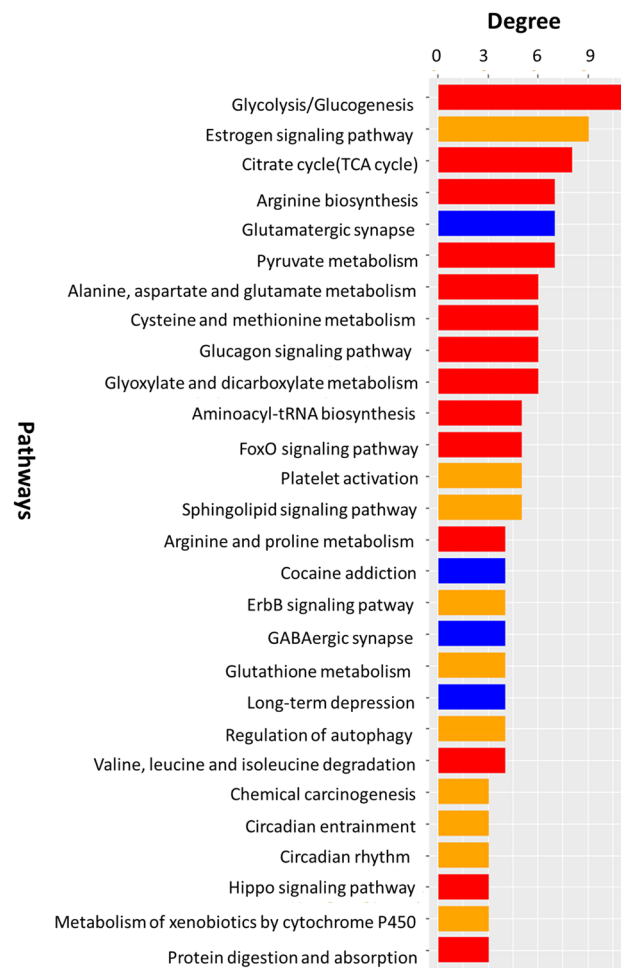
The area under the curve (AUC) values corresponding to different densities are shown in Fig. 4. The highest AUC value was obtained for density 0.1. Then, we employed the clusters generated using density 0.1 for the pathway analysis. Some clusters included non-MD proteins together with currently known MD proteins.

**Identification of SCZ and BD related pathways.** For pathway analysis, we employed the online resources from the Database for Annotation, Visualization and Integrated Discovery (DAVID)<sup>43</sup>, in which a list of genes is considered to conceive the biological meaning for a gene group. In the previous section, based on the highest AUC value we selected 75 statistically significant clusters with  $FDR < 0.05$ . We collected the Uniprot\_Accession\_IDs for the genes, representing the proteins included in those 75 significant clusters. We entered each of these 75 clusters to DAVID separately, and significantly associated KEGG pathways based on count = 2, EASE score = 0.1 and  $FDR \leq 0.05$  were retrieved. Count is the number of common genes/proteins between an input set and a pathway. EASE is the modified Fisher's Exact  $p$ -value. DAVID Function Annotation tool applies Fisher's exact test and provides a  $p$ -value for each pathway indicating its significance. For each cluster we have chosen the top 3 significant pathways. The significant PPI clusters can be considered as sub-systems relevant to SCZ and BD. We hypothesize that the pathways associated to many significant clusters are more relevant to SCZ and BD. Therefore, we finally made a bipartite graph, as shown in Fig. 5, linking these 75 clusters (one set of nodes) with the union of all significant pathways (another set of nodes) chosen for these 75 clusters. From the bipartite graph, 28 high degree pathways (degree  $\geq 3$ ) were selected as SCZ and BD related pathways.

These selected pathways are: glycolysis/gluconeogenesis, estrogen signaling pathway, citrate cycle (TCA cycle), arginine biosynthesis, glutamatergic synapse, pyruvate metabolism, alanine/aspartate/glutamate metabolism, cysteine and methionine metabolism, glucagon signaling pathway, glyoxylate and dicarboxylate metabolism, aminoacyl-t-RNA biosynthesis, FoxO signaling pathway, platelet activation, sphingolipid signaling pathway, arginine/proline metabolism, cocaine addiction, ErbB signaling pathway, GABAergic synapse, glutathione metabolism, long-term depression, regulation of autophagy, valine/leucine/isoleucine degradation, chemical carcinogenesis, circadian entrainment, circadian rhythm, hippo signaling pathway, metabolism of xenobiotics by cytochrome P450, and protein digestion and absorption. These 28 identified pathways and their respective degrees are depicted in Fig. 6, and according to the KEGG pathway database, they could be classified into three main groups, namely energy systems, stress response and neuron systems.



**Figure 5.** Bipartite graph linking significant clusters and corresponding significant pathways.



**Figure 6.** SCZ and BD related pathways predicted in this study (degree  $\geq 3$ , for 28 pathways). Red, yellow and blue bars are pathways related to energy systems, stress response and neuron systems respectively.



## Discussion

Due to the disparate and inconsistent findings from various biomarker studies in mental disorders (MDs), in the present work we aimed to identify the important pathways, predicted to be underlying SCZ and BD based on metabolite biomarkers. These 28 pathways identified herein are mainly involved in three major systems, namely stress response, energy and neuron systems, and they are all significantly related to SCZ and BD, as supported by several evidences reported in the literature and presented here in a simplified way as follow.

The majority of pathways identified in this study are associated with energy metabolism. In fact, among the identified SCZ and BD related pathways, those presenting the highest degree are the (1) *glycolysis/gluconeogenesis*; (2) *pyruvate* (which is the output of the metabolism of glucose) *metabolism*; (3) *glucagon* (which increases glycogenolysis and gluconeogenesis) *signaling pathway*; (4) *citrate cycle* (TCA cycle) (which is an important aerobic pathway for the final steps of the oxidation of carbohydrates and fatty acids); and (5) *FoxO signaling pathway* (which plays important roles in regulation of gluconeogenesis and glycogenolysis by insulin signaling and which affects many cellular physiological processes such as cell cycle, apoptosis, metabolism and oxidative stress, immune regulation). Interestingly, it is well accepted that patients with MDs often present metabolic vulnerabilities with consequent risk of developing cardiometabolic comorbidities<sup>44</sup>, representing one of the leading causes for premature death and reduced lifespan of patients with SCZ or BD<sup>44,45</sup>. In fact, several authors have suggested the involvement of abnormalities in energy metabolism and brain glucose utilization in the pathophysiology of these psychiatric disorders<sup>44,45</sup>. Moreover, proteomic studies with *post-mortem* brains have also provided evidences for energy metabolism dysfunction in several MDs<sup>46</sup>, as well as Schubert et al. reported an increased glycolysis in SCZ and BD<sup>47</sup>.

Moreover, multiple pieces of evidence have also suggested that brain energy metabolism, mitochondrial functions and redox balance are impaired to various degrees in BD and SCZ, and mitochondrial dysfunction and the resultant metabolic changes leading to oxidative stress may also be important etiological factors in the context of these MDs<sup>48–51</sup>.

Mitochondria have a central role in the energy metabolism, and implication of mitochondrial function alterations in the etiology of SCZ is recognized<sup>52</sup>. Furthermore, it is suggested that mitochondrial respiration is downregulated in depression, and upregulated during mania in BD, whilst in SCZ, the number of mitochondria and mitochondrial respiration are both downregulated<sup>53</sup>. Moreover, the mitochondrial dysfunction in blood platelets of patients with manic episodes was proposed as a ‘trait’ marker of BD<sup>54</sup>. One of our identified pathways is *platelet activation*. At this point, it is also worth mentioning that lithium is also key to a wide range of processes at all levels, from neuroprotection to oxidative stress and energy production<sup>55</sup>. In addition, lithium has unquestionable therapeutic superiority for BD treatment, while it also plays an important role in mitochondrial function, which is improved via its role in phospholipid metabolism and inositol depletion<sup>55</sup>.

Possibly as a response to these abnormalities in metabolism and oxidative stress, several compensatory pathways were also identified in the present study, as for instance, the (1) *glutathione metabolism*, which plays important roles in antioxidant defense and its deficiency contributes to oxidative stress; (2) *sphingolipid signaling pathway* which regulates cellular responses to stress; and (3) *ErbB signaling pathway* which regulates diverse biologic responses, including cell proliferation and survival, and regulation of autophagy that is involved in cell growth, survival, development and death, and linked to neurodegeneration in many other disorders, besides being a stress-induced catabolic process. In fact, the main biological alterations of BD and SCZ pertain to inflammation, oxidative stress, membrane ion channels, metabolic dysfunction and circadian system<sup>56–59</sup>. Interestingly, we have also identified here the circadian entrainment and circadian rhythm as highly important pathways associated with the metabolites found in BD and SCZ.

Energy pathways and synaptic function were also implicated in neuropsychiatric disorders such as SCZ and BD<sup>60</sup>. Among the pathways related to neuron systems found in the present work are the (1) *glutamatergic synapse*, (2) *GABAergic synapses*, (3) *long-term depression (LTD)*, (4) *cocaine addiction*, amongst others (as shown in Fig. 6). These pathways are implicated in synaptic/neuronal differentiation, plasticity, and migration, as well as the activation of metabotropic glutamate receptors in the prefrontal cortex induces LTD and reduces the stress-induced anhedonia and other stress-related behavioral impairments in MDs<sup>61</sup>. The LTD is a type of synaptic plasticity in which the efficacy of signal transmission across a synapse continuously decreases after a certain triggering activity, and this activity-dependent plasticity may also result in a persistent enhancement of synaptic transmission<sup>62</sup>. While the LTD pathway has  $F \geq 35$ , the long-term potentiation pathway has  $F < 20$ , therefore, suggesting together that LTD may be implicated in the cognitive dysfunctions observed in SCZ and BD, although not demonstrated or explored in clinics up to now.

We have also identified pathways related to several amino acids metabolism and synthesis (e.g. alanine/aspartate/glutamate metabolism, cysteine and methionine metabolism, arginine/proline metabolism, arginine biosynthesis etc.) and *protein digestion/absorption* pathway, which are important not only for the protein biosynthesis, but also for the functions interrelated with glucose metabolism, synthesis of neurotransmitters, and production of energy, whilst some of them have also the ability to modulate the inflammatory and immune systems. The association between MDs and inflammation/neuroinflammation has been widely discussed and accepted by many, and the correlation of pro-inflammatory markers with symptoms intensity was also reported in SCZ and BD<sup>63</sup>. Two other identified pathways are *Hippo Signaling pathway* and *metabolism of xenobiotics by cytochrome p450*. Cytochrome P450 enzymes (CYPs) play a crucial role in metabolism of xenobiotics in human brain. Recent advances support role of these enzymes in the pathogenesis of psychiatric and neurodegenerative disorders such as depression, and schizophrenia<sup>64</sup>. Hippo Signaling pathway is known to have involvement in stress-related psychiatric disorders<sup>65</sup>.

Taken together, all presented data are in good agreement with the theoretical framework for metabolic comorbidities of mood disorders in which immune system has been likewise “selfish” due to independent energy

consumption, which may compete with the brain (another high energy-consumer) for glucose, which may explain the various conditions of medical impairment, as the Metabolic Syndrome (MetS), obesity, type 2 diabetes mellitus (T2DM) and immune dysregulation, often reported in neuropsychiatric patients<sup>66</sup>.

In addition, some other unexpected pathways were also identified here, as the *estrogen signaling pathway*. This pathway is frequently associated with the activation of various protein-kinase cascades, and *aminoacyl-t-RNA biosynthesis*, which play a central role in protein biosynthesis. Certainly, these protein-kinase cascades and protein biosynthesis deserve special attention. Further studies of these pathways in the context of BD and/or SCZ may have the power to bring new insights into these MDs.

Lastly, good biomarkers for early diagnosis could allow clinical early diagnosis and possibly more adequate treatment<sup>67</sup>. Therefore, early diagnosis of SCZ and BD would be essential to improve outcomes, as early intervention was found to be beneficial for the patients to prevent the cognitive deficits and disabilities if early and properly treated with appropriate pharmacotherapy<sup>68–70</sup>. However, as most of them could not be replicated, unfortunately, no biomarker has been established for differentiating BD and HCs until present<sup>71</sup>, with exception to our most recently work describing a set of metabolic peripheral biomarkers that allow the differential diagnosis between SCZ and BD<sup>20</sup>. However, the absence of any conclusive evidence to identify these severe mental illnesses at an early stage, based on biomarkers alone, led us to propose the integration of these metabolome data by using system biology approach based on protein–protein interactions for identifying pathways underlying SCZ and BD pathophysiology, as presented here. The possible inclusion of some of these more reliable biomarkers, in a study conducted as presented here, may also further increase the power of bringing even more valuable insights into the actual knowledge of pathways underlying these diseases, possibly contributing to early diagnosis and/or a better clinical management.

## Conclusion

In this work we presented a method for identifying pathways underlying Schizophrenia and Bipolar Disorder based on potential biomarkers and PPI network. This study started by collecting a list of potential biomarkers and only those genes that have strong links with the biomarkers and that are connected via reliable PPIs were involved. After constructing a PPI network linked exclusively to SCZ and BD, we identified 75 statistically significant clusters. Based on these clusters, we identified 28 significant pathways, suggesting that SCZ and BD onset may be mainly associated with the abnormality of energy systems, and neuron functions and stress response, which were shown by many to be affected in these MDs. Our results also support the mitochondrial hypotheses for these mental disorders (MDs), and further studies targeting mitochondria function and long-term depression (LTD) may have the power to strongly contribute for the understanding of the mechanism underlying these MDs. Novel pathways, never associated with MDs previously, were also identified here, such as the protein-kinase cascades, LTD and protein biosynthesis. These pathways certainly deserve further attention and studies, aiming to aid in the diagnosis and/or clinical management of MDs. More importantly, we highlight that the present proposed method for finding disease pathways, starting from metabolite biomarkers, is potentially applicable for any other disease.

## Data availability

Dataset may be sent upon request.

Received: 28 December 2020; Accepted: 28 June 2021

Published online: 14 July 2021

## References

- Owen, M. J., Sawa, A. & Mortensen, P. B. Schizophrenia. *Lancet* **388**(10039), 86–97. [https://doi.org/10.1016/S0140-6736\(15\)01121-6](https://doi.org/10.1016/S0140-6736(15)01121-6) (2016).
- Grande, I., Berk, M., Birmaher, B. & Vieta, E. Bipolar disorder. *Lancet* **387**(10027), 1561–1572. [https://doi.org/10.1016/S0140-6736\(15\)00241-X](https://doi.org/10.1016/S0140-6736(15)00241-X) (2016).
- Merikangas, K. R. *et al.* Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch. Gen. Psychiatry* **68**(3), 241–251. <https://doi.org/10.1001/archgenpsychiatry.2011.12> (2011).
- Maggioni, E. *et al.* Common and distinct structural features of schizophrenia and bipolar disorder: The European Network on Psychosis, affective disorders and cognitive trajectory (ENPACT) study. *PLoS ONE* **12**(11), e0188000. <https://doi.org/10.1371/journal.pone.0188000> (2017).
- Johansson, V. *et al.* The schizophrenia and bipolar twin study in Sweden (STAR). *Schizophr. Res.* **204**, 183–192. <https://doi.org/10.1016/j.schres.2018.08.001> (2018).
- da Silva, A. F., Figue, M., van Amelsvoort, T., Veltman, D. & de Haan, L. The revised dopamine hypothesis of schizophrenia: Evidence from pharmacological MRI studies with atypical antipsychotic medication. *Psychopharmacol. Bull.* **41**(1), 121–132 (2008).
- Pogarell, O. *et al.* Dopaminergic neurotransmission in patients with schizophrenia in relation to positive and negative symptoms. *Pharmacopsychiatry* **45**, 36–41. <https://doi.org/10.1055/s-0032-1306313> (2012).
- Ramaker, R. C. *et al.* Post-mortem molecular profiling of three psychiatric disorders. *Genome Med.* **9**(1), 72. <https://doi.org/10.1186/s13073-017-0458-5> (2017).
- Tasic, L. *et al.* Metabolomics and lipidomics analyses by <sup>1</sup>H nuclear magnetic resonance of schizophrenia patient serum reveal potential peripheral biomarkers for diagnosis. *Schizophr. Res.* **185**, 182–189. <https://doi.org/10.1016/j.schres.2016.12.024> (2017).
- Liu, M. L. *et al.* GC-MS based metabolomics identification of possible novel biomarkers for schizophrenia in peripheral blood mononuclear cells. *Mol. BioSyst.* **10**(9), 2398–2406. <https://doi.org/10.1039/c4mb00157e> (2014).
- Nucifora, F. C. Jr., Woznica, E., Lee, B. J., Cascella, N. & Sawa, A. Treatment resistant schizophrenia: Clinical, biological, and therapeutic perspectives. *Neurobiol. Dis.* **131**, 104257. <https://doi.org/10.1016/j.nbd.2018.08.016> (2019).
- Divac, N., Prostran, M., Jakovcevski, I. & Cerovac, N. Second-generation antipsychotics and extrapyramidal adverse effects. *BioMed. Res. Int.* **2014**, 1–6. <https://doi.org/10.1155/2014/656370> (2014).
- Kang, M. G. *et al.* Lithium vs valproate in the maintenance treatment of bipolar I disorder: A post-hoc analysis of a randomized double-blind placebo-controlled trial. *Aust. N. Z. J. Psychiatry* **54**(3), 298–307. <https://doi.org/10.1177/0004867419894067> (2020).



14. Prata, D. P., Costa-Neves, B., Cosme, G. & Vassos, E. Unravelling the genetic basis of schizophrenia and bipolar disorder with GWAS: A systematic review. *J. Psychiatr. Res.* **114**, 178–207 (2019).
15. Bigdeli, T. B. *et al.* Genome-wide association studies of schizophrenia and bipolar disorder in a diverse cohort of US veterans. *Schizophr. Bull.* **47**, 517. <https://doi.org/10.1093/schbul/sbaa133> (2020).
16. Smeland, O. B. *et al.* Genome-wide analysis reveals extensive genetic overlap between schizophrenia, bipolar disorder, and intelligence. *Mol. Psychiatry* **25**(4), 844–853. <https://doi.org/10.1038/s41380-018-0332-x> (2020).
17. Calafato, M. S. *et al.* Use of schizophrenia and bipolar disorder polygenic risk scores to identify psychotic disorders. *Br. J. Psychiatry* **213**(3), 535–541. <https://doi.org/10.1192/bjp.2018.89> (2018).
18. Santoro, M. L. *et al.* A current snapshot of common genomic variants contribution in psychiatric disorders. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **171**(8), 997–1005. <https://doi.org/10.1002/ajmg.b.32475> (2016).
19. Wendt, F. R., Pathak, G. A., Tylee, D. S., Goswami, A. & Polimanti, R. Heterogeneity and polygenicity in psychiatric disorders: A genome-wide perspective. *Chronic Stress (Thousand Oaks)* **4**, 2470547020924844. <https://doi.org/10.1177/2470547020924844> (2020).
20. Tasic, L. *et al.* Peripheral biomarkers allow differential diagnosis between schizophrenia and bipolar disorder. *J. Psychiatr. Res.* **119**, 67–75. <https://doi.org/10.1016/j.jpsychires.2019.09.009> (2019).
21. He, Y. *et al.* Schizophrenia shows a unique metabolomics signature in plasma. *Transl. Psychiatry* **2**(8), e149. <https://doi.org/10.1038/tp.2012.76> (2012).
22. Pedrini, M. *et al.* Advances and challenges in development of precision psychiatry through clinical metabolomics on mood and psychotic disorders. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **93**, 182–188. <https://doi.org/10.1016/j.pnpbp.2019.03.010> (2019).
23. Holmes, E. *et al.* Metabolic profiling of CSF: Evidence that early intervention may impact on disease progression and outcome in schizophrenia. *PLoS Med.* **3**(8), e327. <https://doi.org/10.1371/journal.pmed.0030327> (2006).
24. Cai, H. L. *et al.* Metabolomic analysis of biochemical changes in the plasma and urine of first-episode neuroleptic-naïve schizophrenia patients after treatment with risperidone. *J. Proteome Res.* **11**(8), 4338–4350. <https://doi.org/10.1021/pr300459d> (2012).
25. Sussulini, A. *et al.* Metabolic profiling of human blood serum from treated patients with bipolar disorder employing <sup>1</sup>H NMR spectroscopy and chemometrics. *Anal. Chem.* **81**(23), 9755–9763. <https://doi.org/10.1021/ac901502j> (2009).
26. Burghardt, K., Evans, S., Wiese, K. & Ellingrod, V. An untargeted metabolomics analysis of antipsychotic use in bipolar disorder. *Clin. Transl. Sci.* **8**(5), 432–440. <https://doi.org/10.1111/cts.12324> (2015).
27. Costa, T. B. C. *et al.* Insights into the effects of crack abuse on the human metabolome using a NMR approach. *J. Proteome Res.* **18**(1), 341–348. <https://doi.org/10.1021/acs.jproteome.8b00646> (2019).
28. Sethi, S. *et al.* 1H-NMR, 1H-NMR T2-edited, and 2D-NMR in bipolar disorder metabolic profiling. *Int. J. Bipolar Disord.* **5**(1), 23. <https://doi.org/10.1186/s40345-017-0088-2> (2017).
29. Afendi, F. M. *et al.* KNApSACk family databases: Integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol.* **53**(2), e1. <https://doi.org/10.1093/pcp/pcr165> (2012).
30. Shinbo, Y. *et al.* KNApSACk: A comprehensive species-metabolite relationship database. In *Plant Metabolomics* (eds Saito, K. *et al.*) 165–181 (Springer, 2006).
31. Hossain, S. F. *et al.* Development of a biomarker database toward performing disease classification and finding disease interrelations. *Database* <https://doi.org/10.1093/database/baab011> (2021).
32. Carey, V. *hmdbQuery: Utilities for Exploration of Human Metabolome Database. R Package Version 1.12.0* (2021).
33. Wishart, D. S. *et al.* HMDB 4.0—The human metabolome database for 2018. *Nucleic Acids Res.* **46**, D608–D617 (2018).
34. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
35. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**(11), 2498–2504. <https://doi.org/10.1101/gr.1239303> (2003).
36. Barabasi, A. L. & Albert, R. Emergence of scaling in random networks. *Science* **286**(5439), 509–512. <https://doi.org/10.1126/science.286.5439.509> (1999).
37. Amaral, L. A., Scala, A., Barthelemy, M. & Stanley, H. E. Classes of small-world networks. *Proc. Natl. Acad. Sci. U.S.A.* **97**(21), 11149–11152. <https://doi.org/10.1073/pnas.200327197> (2000).
38. Altaf-Ul-Amin, M., Wada, M. & Kanaya, S. Partitioning a PPI network into overlapping modules constrained by high-density and periphery tracking. *ISRN Biomath.* <https://doi.org/10.5402/2012/726429> (2012).
39. Karim, M. B., Wakamatsu, N. & Altaf-Ul-Amin, M. D. Dedicated to Prof. T. Okada and Prof. T. Nishioka: Data science in chemistry DPCLUSOST: A software tool for general purpose graph clustering. *J. Comput. Aided Mol. Des.* **18**, 76–93. <https://doi.org/10.2751/jcac.18.76> (2017).
40. Altaf-Ul-Amin, M., Shinbo, Y., Mihara, K., Kurokawa, K. & Kanaya, S. Development and implementation of an algorithm for detection of protein complexes in large interaction networks. *BMC Bioinform.* **7**(1), 207. <https://doi.org/10.1186/1471-2105-7-207> (2006).
41. Altaf-Ul-Amin, M. *et al.* DPCLUS: A density-periphery based graph clustering software mainly focused on detection of protein complexes in interaction networks. *JCAC* **7**, 150–156. <https://doi.org/10.2751/jcac.7.150> (2006).
42. Eguchi, R., Karim, M. B., Hu, P., Sato, T. & Ono, N. An integrative network-based approach to identify novel disease genes and pathways: A case study in the context of inflammatory bowel disease. *BMC Bioinform.* **19**(1), 264. <https://doi.org/10.1186/s12859-018-2251-x> (2018).
43. Frank, E. *et al.* Platform for systems medicine research and diagnostic applications in psychotic disorders—The METSY project. *Eur. Psychiatry* **50**, 40–46. <https://doi.org/10.1016/j.eurpsy.2017.12.001> (2018).
44. Correll, C. U. *et al.* Cardiometabolic comorbidities, readmission, and costs in schizophrenia and bipolar disorder: a real-world analysis. *Ann. Gen. Psychiatry* **16**, 9. <https://doi.org/10.1186/s12991-017-0133-7> (2017).
45. Lomholt, L. H. *et al.* Mortality rate trends in patients diagnosed with schizophrenia or bipolar disorder: A nationwide study with 20 years of follow-up. *Int. J. Bipolar Disord.* **7**(1), 6. <https://doi.org/10.1186/s40345-018-0140-x> (2019).
46. Zuccoli, G. S., Saia-Cereda, V. M., Nascimento, J. M. & Martins-de-Souza, D. The energy metabolism dysfunction in psychiatric disorders postmortem brains: Focus on proteomic evidence. *Front. Neurosci.* **11**, 493. <https://doi.org/10.3389/fnins.2017.00493> (2017).
47. Schubert, K. O., Föcking, M. & Cotter, D. R. Proteomic pathway analysis of the hippocampus in schizophrenia and bipolar affective disorder implicates 14–3-3 signaling, aryl hydrocarbon receptor signaling, and glucose metabolism: Potential roles in GABAergic interneuron pathology. *Schizophr. Res.* **167**(1–3), 64–72. <https://doi.org/10.1016/j.schres.2015.02.002> (2015).
48. Cikankova, T. *et al.* Mitochondrial dysfunctions in bipolar disorder: Effect of the disease and pharmacotherapy. *CNS Neurol. Disord. Drug Targets* **16**(2), 176–186. <https://doi.org/10.2174/1871527315666161213110518> (2017).
49. Sampaio, L. R. L. *et al.* Advantages of the alpha-lipoic acid association with chlorpromazine in a model of schizophrenia induced by ketamine in rats: Behavioral and oxidative stress evidences. *Neuroscience* **373**, 72–81. <https://doi.org/10.1016/j.neuroscience.2018.01.008> (2018).
50. Iwata, K. Mitochondrial involvement in mental disorders: Energy metabolism and genetic and environmental factors. *Adv. Exp. Med. Biol.* **1118**, 63–70. [https://doi.org/10.1007/978-3-030-05542-4\\_3](https://doi.org/10.1007/978-3-030-05542-4_3) (2019).
51. Kim, Y. *et al.* Mitochondria, metabolism, and redox mechanisms in psychiatric disorders. *Antioxid. Redox Signal* **31**(4), 275–317. <https://doi.org/10.1089/ars.2018.7606> (2019).

52. Da Silva, T. *et al.* Mitochondrial function in individuals at clinical high risk for psychosis. *Sci. Rep.* **8**(1), 6216. <https://doi.org/10.1038/s41598-018-24355-6> (2018).
53. Morris, G. *et al.* A model of the mitochondrial basis of bipolar disorder. *Neurosci. Biobehav. Rev.* **74**(Pt A), 1–20. <https://doi.org/10.1016/j.neubiorev.2017.01.014> (2017).
54. Hroudová, J. *et al.* Mitochondrial dysfunction in blood platelets of patients with manic episode of bipolar disorder. *CNS Neurol. Disord. Drug Targets* **18**(3), 222–231. <https://doi.org/10.2174/1871527318666181224130011> (2019).
55. Roux, M. & Dosseto, A. From direct to indirect lithium targets: A comprehensive review of omics data. *Metallomics* **9**(10), 1326–1351. <https://doi.org/10.1039/c7mt00203c> (2017).
56. Barandas, R., Landgraf, D., McCarthy, M. J. & Welsh, D. K. Circadian clocks as modulators of metabolic comorbidity in psychiatric disorders. *Curr. Psychiatry Rep.* **17**(12), 98. <https://doi.org/10.1007/s11920-015-0637-2> (2015).
57. Luca, A., Calandra, C. & Luca, M. Gsk3 signalling and redox status in bipolar disorder: Evidence from lithium efficacy. *Oxid. Med. Cell Longev.* **2016**, 3030547. <https://doi.org/10.1155/2016/3030547> (2016).
58. Marco, E. M., Velarde, E., Llorente, R. & Laviola, G. Disrupted circadian rhythm as a common player in developmental models of neuropsychiatric disorders. *Curr. Top. Behav. Neurosci.* **29**, 155–181. [https://doi.org/10.1007/7854\\_2015\\_419](https://doi.org/10.1007/7854_2015_419) (2016).
59. Fond, G., Lançon, C., Korchia, T., Auquier, P. & Boyer, L. The role of inflammation in the treatment of schizophrenia. *Front. Psychiatry* **11**, 160. <https://doi.org/10.3389/fpsy.2020.00160> (2020).
60. Föcking, M. *et al.* Proteomic analysis of the postsynaptic density implicates synaptic function and energy pathways in bipolar disorder. *Transl. Psychiatry* **6**(11), e959. <https://doi.org/10.1038/tp.2016.224> (2016).
61. Joffe, M. E., Santiago, C. I., Engers, J. L., Lindsley, C. W. & Conn, P. J. Metabotropic glutamate receptor subtype 3 gates acute stress-induced dysregulation of amygdalo-cortical function. *Mol. Psychiatry* **24**(6), 916–927. <https://doi.org/10.1038/s41380-017-0015-z> (2019).
62. Massey, P. V. & Bashir, Z. I. Long-term depression: Multiple forms and implications for brain function. *Trends Neurosci.* **30**(4), 176–184. <https://doi.org/10.1016/j.tins.2007.02.005> (2007).
63. Almeida, P. G., Nani, J. V., Oses, J. P., Brietzke, E. & Hayashi, M. A. (2019) Neuroinflammation and glial cell activation in mental disorders. *Brain Behav. Immun. Health* **2**, 100034. <https://doi.org/10.1016/j.bbih.2019.100034> (2020).
64. Duthel, F., Beaune, P. & Lorient, M. A. Xenobiotic metabolizing enzymes in the central nervous system: Contribution of cytochrome P450 enzymes in normal and pathological human brain. *Biochimie* **90**(3), 426–436 (2008).
65. Stepan, J., Anderzhanova, E. & Gassen, N. C. Hippo signaling: Emerging pathway in stress-related psychiatric disorders?. *Front. Psychiatry* **9**, 715 (2018).
66. Buoli, M. *et al.* Biological aspects and candidate biomarkers for psychotic bipolar disorder: A systematic review. *Psychiatry Clin. Neurosci.* **70**(6), 227–244. <https://doi.org/10.1111/pcn.12386> (2016).
67. Vieta, E. & Phillips, M. L. Deconstructing bipolar disorder: A critical review of its diagnostic validity and a proposal for DSM-V and ICD-11. *Schizophr. Bull.* **33**(4), 886–892. <https://doi.org/10.1093/schbul/sbm057> (2007).
68. Lewandowski, K. E. Cognitive remediation for the treatment of cognitive dysfunction in the early course of psychosis. *Harv. Rev. Psychiatry* **24**(2), 164–172. <https://doi.org/10.1097/HRP.000000000000108> (2016).
69. Sanchez-Moreno, J., Martinez-Aran, A. & Vieta, E. treatment of functional impairment in patients with bipolar disorder. *Curr. Psychiatry Rep.* **19**(1), 3. <https://doi.org/10.1007/s11920-017-0752-3> (2017).
70. Murru, A. & Carpinello, B. Duration of untreated illness as a key to early intervention in schizophrenia: A review. *Neurosci. Lett.* **669**, 59–67. <https://doi.org/10.1016/j.neulet.2016.10.003> (2018).
71. Yamagata, A. S. *et al.* Selfish brain and selfish immune system interplay: A theoretical framework for metabolic comorbidities of mood disorders. *Neurosci. Biobehav. Rev.* **72**, 43–49. <https://doi.org/10.1016/j.neubiorev.2016.11.010> (2017).

## Acknowledgements

We are grateful for the technical assistance of Marcela B. Nering, and for the administrative assistance of Rosemary Oliveira.

## Author contributions

All authors equally contributed to this work.

## Funding

This work was supported by the Ministry of Education, Culture, Sports, Science, and Technology of Japan (20K12043 and 16K07223) and NAIST Big Data and Interdisciplinary Project and was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research funded by Japan Agency for Medical Research and Development and the National Bioscience Database Center in Japan. JVN (2019/09207-3) and LCP (2018/03102-2) are recipients of fellowships from FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo). MAFH activities are also supported by FAPESP (2019/13112-8 and 2020/01107-7), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). It was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brazil (CAPES)—Finance Code 001.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to M.A.-U. or M.A.F.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021