Schizophrenia Bulletin Open
doi:10.1093/schizbullopen/sgaa032

Patient Stratification Using Metabolomics to Address the Heterogeneity of Psychosis

Kellie MacDonald[1], Yuting Jiang[1], Ankur Krishnan[1], Sameer Sardaar[1], Bill Qi[1], Aristotelis Eleftheriadis[1], Stephen J. Glatt[2], Ridha Joober[3], John Mitchell[1,4], Karim Tabbane[3], and Yannis Trakadis*[,1,4]

[1]Department of Human Genetics, Faculty of Medicine, McGill University, Montreal, Canada; [2]Psychiatric Genetic Epidemiology & Neurobiology Laboratory (PsychGENe Lab), SUNY Upstate Medical University, Syracuse, NY; [3]Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, Canada; [4]Department of Medical Genetics, McGill University Health Center, Montreal, Canada

*To whom correspondence should be addressed; Yannis J. Trakadis, Department of Medical Genetics, McGill University Health Centre, Room A04.3140, Montreal Children's Hospital, 1001 Boul. Décarie, Montreal, Quebec, Canada, H4A 3J1; tel: 514-412-4427, fax: 514-412-4296, e-mail: yannis.trakadis@mcgill.ca

Psychosis is a symptomatic endpoint with many causes, complicating its pathophysiological characterization and treatment. Our study applies unsupervised clustering techniques to analyze metabolomic data, acquired using 2 different tandem mass spectrometry (MS-MS) methods, from an unselected group of 120 patients with psychosis. We performed an independent analysis of each of the 2 datasets generated, by both hierarchical clustering and k-means. This led to the identification of biochemically distinct groups of patients while reducing the potential biases from any single clustering method or datatype. Using our newly developed robust clustering method, which is based on patients consistently grouped together through different methods and datasets, a total of 20 clusters were ascertained and 78 patients (or 65% of the original cohort) were placed into these robust clusters. Medication exposure was not associated with cluster formation in our study. We highlighted metabolites that constitute nodes (cluster-specific metabolites) vs hubs (metabolites in a central, shared, pathway) for psychosis. For example, 4 *recurring* metabolites (spermine, C0, C2, and PC.aa.C38.6) were discovered to be significant in at least 8 clusters, which were identified by at least 3 different clustering approaches. Given these metabolites were affected across multiple biochemically different patient subgroups, they are expected to be important in the overall pathophysiology of psychosis. We demonstrate how knowledge about such hubs can lead to novel antipsychotic medications. Such pathways, and thus drug targets, would not have been possible to identify without patient stratification, as they are not shared by all patients, due to the heterogeneity of psychosis.

Key words: machine learning/schizophrenia/omics/clustering

Introduction

The heterogeneity of psychiatric conditions renders their pathophysiological characterization and treatment-optimization rather challenging. For example, previous studies have considered schizophrenia (SCZ) to be a common symptomatic endpoint resulting from the combination of a variety of brain dysfunctions, suggesting that many patients with SCZ are unlikely to possess the same disease etiopathology.[1,2] The difficulty to predict how SCZ patients will respond to certain antipsychotic treatments underscores this heterogeneity, as up to 30% of patients with SCZ do not respond to conventional antipsychotic medications.[3] Additionally, given that the current diagnostic criteria, outlined by the Diagnostic and Statistical Manual of Mental Disorders fifth edition (DSM-5),[4] rely mainly on the patient's self-reported symptoms, adequate identification of homogeneous groups of patients at the pathophysiology level is not possible. This, in turn, hinders the elucidation of the different underlying pathways affected and the identification of novel candidate drugs for patients with psychiatric disease.[5]

Metabolomics focuses on the identification and quantification of the small molecules, called metabolites, produced from metabolic pathways in the body. It relies on high-throughput technologies, such as mass spectrometry (MS), which are capable of rapidly

surveying hundreds of metabolites at once. Given that metabolites are the products of metabolism, they are the downstream functional markers of combinations of complex environmental, genomic, and proteomic changes in an individual. Metabolomics can, therefore, be used to study the chemical composition of different body specimens, such as blood, urine, or cerebrospinal fluid (CSF), aiming to identify specific sets of chemical changes characteristic of any given illness. While the use of metabolomics is becoming more mainstream in other medical fields, only a few metabolomic studies have been completed in patients with SCZ or psychosis. All such published studies focus on biomarker discovery for SCZ or psychosis rather than on patient stratification.[6,7]

This manuscript takes advantage of unsupervised clustering techniques applied to metabolomics data for the stratification of patients with psychosis. Our hypothesis is that this approach can provide the basis for patient stratification leading to a better understanding of the pathophysiology of psychosis. To this end, we analyzed urine samples in a group of unselected patients with psychosis. Urine was the biological fluid of choice in this study to facilitate clinical translation. Urine is the least invasive sample that can be collected, is easy to store and requires limited preparation for metabolomic analysis. Of note, no controls were required in this study, since our goal was not to identify biomarkers for affected vs unaffected status, but rather to stratify patients based on their metabolic profiles, and to characterize the different clusters.

Currently, in other medical fields, stratifying patients with a common diagnosis into subgroups, also known as endotypes, through the use of metabolomic data and clustering methods is becoming more routine in the research of other heterogeneous conditions such as breast cancer, asthma, and Alzheimer's.[8–10] However, to our knowledge, no similar studies have been completed in psychosis or SCZ.

What is needed clinically is not a test for the presence or absence of psychosis; this is usually quite obvious to the clinician. In our opinion, what is really needed, for both clinical and research purposes, is a basis for distinguishing types of psychosis. Knowledge of the differences and similarities between such distinct groups of patients could advance our understanding of the heterogeneous concept of psychosis and potentially lead to novel antipsychotic medications for patients with a specific type of psychosis, or for psychosis in general.

Methods

Participant Recruitment

The study protocol was approved by the Research Ethics Board of the Douglas Mental Health University Institute (DMHUI). One hundred twenty consecutive, unselected patients with a history of psychosis, between the ages of 20 and 75, were recruited from the outpatient and inpatient services at the DMHUI. As per our inclusion criteria, adults with a clinical diagnosis of psychosis, as confirmed by a psychiatrist in accordance with the DSM[4] and the Positive and Negative Syndrome Scale (PANSS),[11] were included in this study. Patients who had not been evaluated by a psychiatrist, patients whose clinical history was not available (ie, cases, where neither the patient, nor the caregiver was capable of providing the patient's medical history and the information could not be derived from the patient's chart), and patients who did not speak English or French, were excluded from this study.

Study Procedures

Consenting patients were asked to complete a self-report clinical questionnaire, designed based on the clinical features of inborn errors of metabolism (IEM) known to be associated with psychosis.[12] The self-report questionnaire inquired about the patient's family history, past medical history, and current physical symptoms. Study participants also underwent a physical examination, which evaluated the patients for pertinent neurological abnormalities, as well as abnormalities in the patients' abdomen, skin, muscles, and skeleton.[12] A trained research assistant helped participants with the questionnaire and performed the physical exam.

Approximately 10–15 ml of urine was then collected from each participant and kept on ice until frozen at −80°C. The urine samples were shipped on ice to The Metabolomics Innovation Centre (TMIC) at the University of Alberta in Edmonton, Canada, where they underwent a targeted analysis on the DI/LC-MS/MS (direct injection/liquid chromatography tandem mass spectrometry) chemical analysis platform. The targeted analysis involved the measurement of a specified list of metabolites, as already established by the TMIC.[13] This list of metabolites was compiled and validated by TMIC, a leading institution in the field of metabolomics, and consists of all metabolites that can be detected and quantified in urine in humans. A recent systematic review on metabolomics of psychosis provides support that the metabolites targeted in our current study capture the most important metabolites identified as biomarkers in previous studies for psychosis.[7] For each patient, 2 different data sets were generated, one using LC-MS/MS and the other based on DI-MS/MS analysis. DI-MS/MS works by directly injecting the samples into the mass spectrometer for analysis, while LC-MS/MS works by running the samples through a column that helps to separate the molecules of interest using various chemical properties (eg, polarity, ionization capabilities) and solvents to extract the molecules. Thus, the metabolites identified by each of these techniques are different. Two

different unsupervised clustering methods were then used to analyze each of these data sets, as described below.

Unsupervised Clustering Methods

The raw metabolite values of the 38 metabolites targeted by LC-MS/MS and the 109 metabolites targeted by DI-MS/MS were first standardized using a z-score approach. Calculating the z-score for a given metabolite of a specific patient involved taking the patient's value for the metabolite targeted, subtracting it from the mean of that same metabolite across all the patients, then dividing the result by the standard deviation of that same metabolite across all the patient values. The z-score itself is representative of how many standard deviations away from the mean the value of the metabolite in question is for a given patient.

The standardized values of the metabolites across all patients were then used to perform clustering analysis using 2 well-recognized techniques for unsupervised clustering: k-means and hierarchical clustering (HC). For each technique, patients were placed into 1 of 10 clusters based on the similarity of their metabolite profiles. The optimum number of clusters for the k-means analysis was determined based on the Elbow method,[14] and consequently, a similar number of clusters were then identified for the HC analysis. We also developed an additional method for subgrouping patients, which is based on converging clustering of patients (namely, the "robust clustering approach").

Robust Clustering Method

Our "Robust Clustering" method identifies the patients clustering together based on all 4 combinations of the different methods from the ones described above. The robust clustering method consisted of comparing the cluster that each patient was placed in for the (1) HC and (2) k-means analysis. Given that we analyzed each data set (LC-MS/MS vs DI-MS/MS) separately, for patients to be grouped in the same robust cluster, they needed to have been grouped into the same clusters based on 4 approaches (the HC (1a) and k-means (2a) analysis of the DI-MS/MS data, as well as, the HC (1b) and k-means (2b) analysis of the LC-MS/MS data). Patients grouped into the same clusters for each of these 4 clustering approaches constituted the new, robust clusters (supplementary table S1).

Permutation Analysis of Robust Clustering Results. We performed a permutation analysis to verify the statistical significance of the overlaps observed in the clustering results from the different datasets (LC-MS/MS vs DI-MS/MS). Specifically, the adjusted Rand score (ARS) between the clustering assigned to each individual from k-means clustering of the LC-MS/MS and DI-MS/MS datasets was calculated. Next, to calculate the statistical significance of obtaining the observed result, we calculated the ARS between the k-means clustering of LC-MS/MS dataset and a random clustering result generated by permuting the clustering result from k-means clustering of DI-MS/MS data. We repeated the permutation procedure for 100 000 iterations to derive the distribution of ARS based on random clustering. The *P*-value was then defined as the percentage of ARS values that were higher than the observed result. The standard *P*-value cutoff of .05 was used to denote significance.

The above procedure was also repeated for the HC clustering of LC-MS/MS and DI-MS/MS datasets.

Consensus Clustering Method

Consensus clustering is an unsupervised clustering method that uses algorithms to sample a dataset through multiple iterations, in order to represent the consensus of any discovered clusters, providing information on the number of clusters and the membership within them.[15] Using the ConsensusClusteringPlus package in R[16] and the z-scores calculated from the raw metabolomic data (see Unsupervised Clustering Methods for greater detail on this calculation), the consensus clustering was performed (agglomerative HC over 1000 iterations, with each iteration randomly sub-sampling 80% of the cohort). Euclidean distance measure was selected, along with Ward for both innerLinkage and finalLinkage. Using the Elbow method,[14] k was then chosen to be 6, based on the data from the change in area under the cumulative distribution function (CDF) graphical output.

Characterization of the Clusters

Once the patients had been successfully stratified into groups, different strategies were implemented to further characterize each cluster, as described below.

Biochemical Characterization of the Clusters. To biochemically characterize a cluster, we used an approach that we termed the ratio-of-means approach. The ratio-of-means approach is centered on the idea that the metabolites with values falling more than 2 SDs away from the mean would be expected to contribute to the clustering effect (supplementary figure S1).

It is important to highlight that we are comparing the value of a given metabolite in a cluster to the value of that metabolite corresponding to the remaining patients in our cohort. Our focus is on determining what makes each cluster distinct, rather than on identifying biomarkers for diagnosing patients with psychosis from controls (supplementary methods).

Pharmacological and Clinical Features Characterization of the Clusters. To explore if any medications might

be influencing the clustering results, a Fisher's exact test was performed to determine whether the ratio of patients taking a medication in a cluster was statistically different from patients outside of a cluster. Furthermore, the Fisher's exact test was also used for clinical features to determine whether any clusters were associated with specific features. Lastly, as age can be a contributing factor in regard to metabolomic changes, *t*-tests with FDR adjustment were performed for patient age of each cluster compared to the age outside the cluster.

Pathway and Candidate Drug Analysis. Using the "Pathway analysis feature" in MetaboAnalyst 4.0,[17] a biochemical pathway analysis was conducted for all the significant metabolites identified from the different clusters. The pathways highlighted by the different clustering approaches were identified. Our hypothesis was that if a pathway was highlighted in multiple patient clusters, it must play a more central role in the pathophysiology of psychosis. Thus, prior evidence for a link with psychosis was explored for pathways identified by at least 3 different clustering approaches and identified to be significant for at least 15 different clusters. DrugBank 5.0,[18] a database containing biochemical and pharmacological information on various drugs, was searched for candidate drugs targeting these pathways.

Results

Patient Demographics

The demographic information of our patients, along with the medical findings, are summarized in table 1. All patients, with the exception of one, were receiving antipsychotic medication at the time of urine collection (supplementary table S2).

Unsupervised Clustering Outcomes

Ten clusters were obtained from the k-means and HC analysis of our 2 datasets (see Unsupervised Clustering Methods), capturing all patients. Using our newly developed robust clustering method, which is based on patients consistently grouped together through different methods and datasets, a total of 20 clusters were ascertained and

Table 1. Demographics and Clinical Features of Patients With Psychosis

Psychosis Group (N = 120)			
	Mean (SD)		N (%)
Age (y)	42.9 (11.45)	Neurological features	
	N (%)	School problems	69 (58)
Demographics		Progressive neurological deterioration	8 (7)
Male	79 (66)	Movement difficulties	33 (28)
Female	38 (32)	Paralysis	10 (8)
Undisclosed	3 (3)	Sensory problems	12 (10)
Family history		Confusion	36 (30)
Consanguineous parents	3 (3)	Neuropathy	45 (38)
Early death	24 (20)	Seizures	11 (9)
Recurrent pregnancy loss	9 (8)	Stroke	8 (7)
Psychiatric illness	81 (68)	Balance problems	31 (26)
Heart problems	16 (13)	Cardiovascular features	
		Tachycardia	22 (18)
		Hypertension	23 (19)
		Heart problems	8 (7)
		Dietary and gastrointestinal features	
		Food aversion	19 (16)
		Constipation	26 (22)
		Diarrhea	14 (12)
		Abdominal cramps	18 (15)
		Vomiting	20 (17)
		Catabolic crises	12 (10)
		Other clinical features	
		Kidney problems	10 (8)
		Dark urine (possible rhabdomyolysis)	4 (3)
		Recurrent infections	15 (13)
		Skin problems	31 (26)
		Diabetes	24 (20)
		Eye problems	85 (71)

Note: *N*, Number, %, Percentage.

Table 2. The Metabolites That Were Discovered to be Significant From the Ratio-of-Means Analyses of all Clusters

Metabolite Name	Number of Clusters That Identified a Metabolite as Being Significant for a Given Clustering Approach					Total Number of Clusters That Identified the Metabolite
	HC of LC-MS/MS	k-Means of LC-MS/MS	HC of DI-MS/MS	k-Means of DI-MS/MS	Robust Clustering by k-Means and HC	
4-Hydroxyproline	2	1	-	-	1	4
Carnosine	1	-	-	-	3	4
Citrulline	-	1	-	-	3	4
Glutamic Acid	1	2	-	-	4	7
Proline	1	2	-	-	-	3
Spermidine	4	4	-	-	-	8
Spermine[a]	2	2	-	-	4	8
C0[a]	-	-	3	1	4	8
C14.1	-	-	2	2	-	4
C2[a]	-	-	2	2	4	8
C3	-	-	2	1	3	6
lysoPC.a.C20.4	-	-	1	-	2	3
PC.aa.C34.1	-	-	1	1	1	3
PC.aa.C34.2	-	-	2	1	1	4
PC.aa.C34.3	-	-	2	1	2	5
PC.aa.C36.2	-	-	3	2	2	7
PC.aa.C36.3	-	-	2	2	2	6
PC.aa.C36.4	-	-	4	2	1	7
PC.aa.C38.3	-	-	2	2	1	5
PC.aa.C38.4	-	-	4	2	1	7
PC.aa.C38.5	-	-	3	2	2	7
PC.aa.C38.6[a]	-	-	4	2	2	8
PC.ae.C30.1	-	-	-	4	4	8
SM.C20.2	-	-	1	1	3	5
SM.C22.3	-	-	-	-	4	4
Hexose	-	-	1	3	2	6

Note: HC, Hierarchical clusters; DOPA, Dihydroxyphenylalanine; PC, Phosphatidylcholine; LysoPC, Lysophosphatidylcholine; SM, Sphingomyelin; SM (OH), Hydroxysphingomyelin; a, acyl; aa, diacyl; ae, alkyl-acyl; LC-MS/MS, Liquid chromatography tandem mass spectrometry; DI-MS/MS, Direct injection tandem mass spectrometry. All metabolites seen in the table above were identified to be significant (2 SDs away from the mean) by the ratio-of-means analysis. Due to the large number of metabolites identified, only metabolites identified in 3 or more clusters are listed in this table. Four metabolites (spermine, C0, C2, and PC.aa.C38.6) were identified by 3 or more clustering approaches and were identified by a total of at least 8 clusters.
[a]Metabolites identified by 3 or more clustering approaches in a total of 8 or more clusters.

78 patients (or 65% of the original cohort) were placed into these robust clusters (supplementary table S2). Based on permutation analysis of the clustering results, we were able to determine that the overlap/similarity between the clustering results was significantly higher than random chance. For k-means, the ARS between the clustering obtained from the DI-MS/MS dataset and the LC-MS/MS dataset was 0.132 (P-value = 0). Correspondingly for HC, the ARS was 0.109 (P-value = .00026).

Following consensus clustering of all the metabolite data, it was determined that 85% of patients, who were grouped together following robust clustering, were also grouped together in the consensus clusters.

The analysis of the medications taken by patients in each of the robust clusters suggested that the medications were not contributing to clustering as none of the p-values from the Fisher's exact test were significant. There was no significance in cluster formation for clinical features or age.

The recurring metabolites that were discovered to be significant from the ratio-of-means analyses of the different clusters are summarized in table 2. Four metabolites (spermine, C0, C2, and PC.aa.C38.6) were highlighted in at least 8 clusters, which were identified by at least 3 different clustering approaches.

Table 3 summarizes the metabolites that were identified as being significant in multiple clusters and highlights the corresponding clustering method.

Pathway Analysis Findings for Significant Metabolites Identified From the Different Clustering Approaches

Table 4 highlights the recurrent pathways, ie, the pathways identified based on different clusters, and approaches, using MetaboAnalyst. Arginine and proline metabolism, beta-alanine metabolism, glutathione metabolism, and glycerophospholipid metabolism

Table 3. Metabolites Identified as Being Significant for Multiple Clusters Within a Clustering Method

Metabolite	HC of LC-MS/MS	k-Means of LC-MS/MS	HC of DI-MS/MS	k-Means of DI-MS/MS	Robust Clustering by k-Means and HC
4-Hydroxyproline	✓	-	-	-	-
Aspartic Acid	-	-	-	-	✓
Carnosine	-	-	-	-	✓
Citrulline	-	-	-	-	✓
DOPA	-	-	-	-	✓
Glutamic Acid	-	✓	-	-	✓
Methionine	-	-	-	-	✓
Proline	-	✓	-	-	-
Putrescine	-	✓	-	-	-
Spermidine	✓	✓	-	-	-
Spermine	✓	✓	-	-	✓
C0	-	-	✓	-	✓
C14.1	-	-	✓	✓	-
C14.1.OH	-	-	✓	-	-
C2	-	-	✓	✓	✓
C3	-	-	✓	-	✓
C5	-	-	-	-	✓
C6.(C4.1.DC)	-	-	-	-	✓
C8	-	-	✓	-	-
C9	-	-	-	-	✓
lysoPC.a.C20.4	-	-	-	-	✓
PC.aa.C30.2	-	-	-	-	✓
PC.aa.C32.3	-	-	-	-	✓
PC.aa.C34.2	-	-	✓	-	-
PC.aa.C34.3	-	-	✓	-	✓
PC.aa.C36.2	-	-	✓	✓	✓
PC.aa.C36.3	-	-	✓	✓	✓
PC.aa.C36.4	-	-	✓	✓	-
PC.aa.C38.3	-	-	✓	✓	-
PC.aa.C38.4	-	-	✓	✓	-
PC.aa.C38.5	-	-	✓	✓	✓
PC.aa.C38.6	-	-	✓	✓	✓
PC.aa.C40.1	-	-	-	-	✓
PC.aa.C40.2	-	-	-	-	✓
PC.aa.C42.4	-	-	-	-	✓
PC.ae.C30.0	-	-	-	-	✓
PC.ae.C30.1	-	-	-	✓	✓
PC.ae.C42.0	-	-	-	-	✓
SM.C20.2	-	-	-	-	✓
SM.C22.3	-	-	-	-	✓
SM.C26.0	-	-	-	-	✓
SM.C26.1	-	-	-	-	✓
Hexose	-	-	-	✓	✓
Total	3	5	15	11	32

Note: HC, Hierarchical clusters; DOPA, Dihydroxyphenylalanine; PC, Phosphatidylcholine; LysoPC, Lysophosphatidylcholine; SM, Sphingomyelin; a, acyl; aa, diacyl; ae, alkyl-acyl; LC-MS/MS, Liquid chromatography tandem mass spectrometry; DI-MS/MS, Direct injection tandem mass spectrometry.

were all pathways that were highlighted by at least 2 different metabolites and were significant for at least 15 different clusters, which were identified by at least 3 different clustering approaches. Comparing the results of the pathway analysis from the consensus clusters revealed all of the aforementioned pathways, as well as arachidonic acid, which was also identified through robust clustering, and tryptophan metabolism (supplementary table S3).

Discussion

Clusters of Patients With Psychosis and Characterization

No correlation was found between the DSM-V based diagnoses and the clusters identified (supplementary table S2). This was expected, given the heterogeneity of each DSM diagnosis. Analyzing each of the 2 datasets generated (LC-MS/MS vs DI-MS/MS dataset), by both

Table 4. Summary of Pathways Identified by 3 or More Clustering Approaches and That Were Significant for At Least 15 Different Clusters

	Number of Clusters That Identified the Pathway for a Given Clustering Approach					
Pathway Name	HC of LC-MS/MS	k-Means of LC-MS/MS	HC of DI-MS/MS	k-Means of DI-MS/MS	Robust Clustering by k-Means and HC	Total Number of Clusters That Identified the Pathway
Alpha-linolenic acid metabolism	-	-	4	4	9	17
Arginine and proline metabolism	8	7	-	-	5	20
Beta-alanine metabolism	5	4	-	-	7	16
Glutathione metabolism	7	7	-	-	3	17
Glycerophospholipid metabolism	-	-	4	4	7	15
Linoleic acid metabolism	-	-	4	4	10	18

Note: HC, Hierarchical clustering; LC-MS/MS, Liquid chromatography tandem mass spectrometry; DI-MS/MS, Direct injection tandem mass spectrometry.

HC and k-means, led to significant overlap of how the patients clustered together. As discussed above, the patients who were consistently grouped together through different clustering methods and datasets formed the robust clusters (supplementary table S1). This method of patient stratification reduces the potential biases from any single clustering method or datatype, as it takes advantage of multiple clustering methods applied to multiple feature spaces (in this case, 2 independent lists of metabolites). 85% of patients grouped together into a robust cluster were also grouped together in a consensus cluster, which increases the confidence of our patient clusters.

When examining the clusters obtained from the robust clustering method, we found that 65% of our original cohort could be placed into a robust cluster. Medication exposure, age, and clinical characteristics were not associated with cluster formation in our study. This finding is compatible with the possibility that each cluster corresponds to a subtype of psychosis with different pathophysiology. Each of these distinct groups of patients could represent a different type of psychosis, secondary to a specific combination of genetic and environmental factors. Given the symptomatic endpoint across the different clusters is the same (namely, psychosis), one would also expect some overlap across the distinct pathophysiology of different clusters (albeit, not across all clusters).

Based on the ratio-of-means approach undertaken to identify significant metabolites for each cluster, several recurrent metabolites (ie, important for different clusters) were identified (table 2). Pathway analysis, of the significant metabolites identified for each cluster of patients, found that lipid, amino acid, and energy metabolism may be altered in patients with psychosis, as supported by our literature review. The 4 most robust metabolites identified in this study were spermine, carnitine (C0), acetylcarnitine (C2), and phosphatidylcholine diacyl C38:6 (PC. aa.C38.6). Some of these metabolites have published evidence for a link with psychosis. For example, spermine is a polyamine that has been shown to target the N-methyl-D-aspartate (NMDA) receptor, dysregulation of which is thought to provoke psychosis.[19] Antipsychotic-naïve first-episode psychosis patients have higher levels of spermine in their serum, and treatment with an antipsychotic drug results in the levels normalizing to a range similar to control subjects.[20,21] It should be highlighted that any biomarkers or significant metabolites that are determined using the methods described herein, are merely reflective of the cluster in question (ie, what makes it distinct from the other clusters). They have been ascertained through comparison to the entire cohort, and we do not claim that they would be useful in the diagnosis of psychosis from those without disease.

Having said this, one of the major advantages of our study is that, through patient stratification, it allows us to identify recurrent metabolites (eg, spermine, carnitine, acetylcarnitine, and phosphatidylcholine diacyl C38:6), affected across multiple biochemically different patient subgroups, and thus deemed to be important in the pathophysiology of psychosis, in general. Due to the heterogeneity at the pathophysiology level, it would not have been possible to identify these biomarkers if the entire heterogeneous population of patients with psychosis was analyzed as a whole. As described in the introduction, not all patients with SCZ have the same pathophysiology and our hypothesis is that by subgrouping patients into biochemically distinct clusters, one can learn more about the metabolites that constitute nodes (cluster-specific metabolites) vs hubs (metabolites in a central, shared, pathway) for SCZ. Identification of hubs, and pathways

of hubs, can improve our understanding of the pathophysiology of psychosis and lead to novel antipsychotics.

Literature Review for the Recurrent Pathways Identified

Pathway analysis of the significant metabolites (table 2) revealed several metabolic pathways which may be implicated in the pathophysiology of psychosis (table 4). As per table 4, 4 pathways were identified by at least 2 metabolites, based on 3 or more clustering approaches, and in more than 15 clusters. These included arginine and proline metabolism, beta-alanine metabolism, glutathione metabolism, and glycerophospholipid metabolism. Our hypothesis is that since these pathways were shown to be important in several different subgroups of patients, they may be more central to the development of SCZ (hubs). The findings from the literature review of these significant pathways support this point and highlight how our approach can potentially lead to the development of novel therapeutics.

There is evidence for Glutathione (GSH) being a primary vulnerability factor in the development of SCZ.[22] Interestingly, spermine (discussed above) has previously been shown to prevent the release of GSH in the mitochondria of rat livers,[23] and may have similar implications in brain mitochondria. Studies have also reported reduced levels of GSH in the brain and blood of early psychosis and chronic SCZ patients.[24,25]

Beta-alanine is synthesized from spermine. It is considered to be a direct precursor of pantothenic acid, a water-soluble B vitamin involved in the synthesis of coenzyme A (CoA)[26] that has implications in the pathogenesis of SCZ.[27] Beta-alanine can also be reversibly converted to carnosine, an antioxidant and antiglycation agent co-localized and released with glutamate, and preliminary research on this di-peptide as adjuvant to traditional therapies for SCZ seems to be effective in treating some of the executive dysfunction symptoms common to SCZ.[28]

Aberrations in arginine and proline metabolism have been implicated in the pathophysiology of psychosis, SCZ and other neuropsychiatric disorders.[29–31] Arginine and proline metabolism also includes the metabolism of putrescine, which is an intermediate in the conversion of arginine to spermidine and ultimately spermine, a precursor in GSH and beta-alanine biosynthesis, as mentioned above.[32] The polyamines, spermine, spermidine, and putrescine are also linked to neurotransmitter systems, and thus SCZ, via their implication in regulating the NMDA receptors.[19,33,34] Studies have shown that the expression of genes involved in the synthesis of polyamines is reduced in the brains of schizophrenic patients.[35] Many studies show an accelerated glycerophospholipid metabolism in drug-naïve schizophrenic patients.[36–38] Human brain development, including synaptic remodeling, learning and memory processing, has been shown to be disturbed by abnormal glycerophospholipid metabolism[39,40] and could help to explain some of the symptomatology that is consistent with SCZ.[41]

Candidate Drugs for Psychosis

In addition to potentially advancing our understanding of the different pathways affected in psychosis, our approach may also provide insight into the development of novel antipsychotics. Searching DrugBank, for drugs targeting the pathways identified across multiple patient clusters, highlighted different candidate drugs for psychosis. Several of these candidates have published evidence supporting their role in psychosis (arachidonic acid,[42] glycine,[43,44] icosapent,[45–55] L-arginine,[55] L-cysteine,[56–59] pyridoxal phosphate,[60] and zinc[61–63]). For example, zinc has targets in the beta-alanine and glutathione metabolic pathways and is a known inhibitor of NMDA receptors. It has been studied widely using animal models and human studies for its antidepressant-like effects,[64,65] and it is thought that zinc deficiency and/or impaired release of zinc from the hippocampus may play an important role in the pathophysiology of the psychosis-like state and in SCZ.[66]

Conclusions

Our study focused on urine samples from patients with psychosis because urine is the least invasive biospecimen, and thus more practical for clinical translation of the proposed approach. The independent application of different tandem mass spectrometry methods and unsupervised clustering techniques led to the identification of biochemically distinct groups of patients while reducing the potential biases from any single clustering method or datatype.

Overall, the pathways identified for each unsupervised clustering approach (k-means and HC) were similar. Moreover, although the metabolites targeted by the DI-MS/MS and LC-MS/MS technologies are different, the 2 methods led to significant overlap of how the patients clustered together: the 20 robust clusters identified included 78 patients (ie, 65% of the entire patient population). Overall, our results indicate that our robust clustering approach constitutes a promising method of clustering patients into biochemically distinct groups. Thus, the approach described herein has the potential to facilitate a pathophysiology-based stratification of patients with psychiatric disease. The clusters of patients identified may represent patients with similar pathophysiology, potentially due to distinct combinations of environmental, genetic, and/or other medical causes.

We highlight some metabolites identified in multiple clusters in our study (namely, spermine, carnitine, acetylcarnitine, and phosphatidylcholine diacyl C38:6), which we conclude may play a central role for psychosis.

Identification of such metabolites can improve our understanding of the pathophysiology of psychosis. As described in the introduction, not all patients with SCZ have the same pathophysiology and our hypothesis is that by subgrouping patients into biochemically distinct clusters, one can learn more about the metabolites that constitute nodes (cluster-specific, or else "psychosis subtype-specific," metabolites) vs hubs (metabolites in a central, shared, pathway of psychosis). We argue that the metabolites shared across several clusters represent metabolite disturbances, which are more common in the pathophysiology of SCZ. Hence, the pathways corresponding to these metabolites may constitute excellent drug targets for the development of novel antipsychotics for patients with SCZ in general. In contrast, pathways highlighted by rare distinct cluster-specific significant metabolites may be important in the development of treatments that would be efficient only in the patients of that cluster (precision medicine).

We demonstrate how such knowledge can lead to novel antipsychotics. Until larger studies are available to help categorize a large proportion of patients with SCZ into specific subtypes, and identify/validate treatments for each of these subtypes (precision medicine), our approach based on stratification can potentially yield new candidate drugs for the overall heterogeneous patient population with SCZ. This would still be following the current "trial and error" paradigm for treatment prioritization but may be helpful by increasing the treatment options available for SCZ, in general.

In conclusion, our study highlights the importance of metabolomics and unsupervised clustering in the stratification of patients with psychosis. It provides new insights about how one can take advantage of these tools to address the heterogeneity, and advance precision medicine, in psychiatric disease.

Limitations and Future Directions

Patients clustering together need to be further characterized prospectively. They may have similar complications or similar response to specific antipsychotics. Finally, more targeted treatments (known or novel), focused on the pathways identified based on the metabolomic data of patients within each cluster, could be made available for them.

As a next step, a larger study with the same approach that focuses on urine, blood, and CSF samples, involving repeated collection at different time intervals for a time-series analysis, would be recommended. Also, given the challenges of using tandem MS[67], the complementary use of proton nuclear magnetic resonance (NMR) would be recommended. The use of untargeted metabolic analysis would be highly encouraged given novel or unexpected metabolites could prove to be important in the stratification process.

Supplementary Material

Supplementary data are available at *Schizophrenia Bulletin Open* online.

Funding

The metabolomic analysis of the urine samples was supported by the Garrod Association and Dr. David Wishart's lab. The funding for the research assistant was supported by Dr. Joober's lab. Dr. Trakadis receives funding for his research from the Montreal Children's Hospital Foundation.

Acknowledgments

We thank Nadia Zeramdini from Dr. Joober's lab for helping to recruit the patients, provide the questionnaires and perform the physical exam, along with the psychiatrists. Dr. David Wishart and his team at The Metabolomic Innovation Centre (Dr. Rosa Vázquez-Fresno, Dr. Rupa Mandal, and Dr. BeomSoo Han) should also be thanked for running the urine samples on their platforms and providing us with the metabolomics data. Y.T. conceived and coordinated the project. A.E. was responsible for preparing the clinical and metabolomics data sets for analysis by other members of the team. B.Q., Y.J., and S.S. were responsible for the bioinformatics and machine learning aspects of this study. They developed the different scripts used for the analysis (eg, ratio-of-means and z-score approaches, robust clustering approach), under the supervision of Y.T. who designed the approaches. The methodology used was verified by our collaborator S.J.G., specializing in genetic epidemiology. S.S. performed the clustering analysis. A.K. determined the significant metabolites and performed the pathway analysis with these metabolites. Y.J. performed all of the same steps on the DI-MS/MS data. A.K., Y.J., and K.M., all participated in the literature review of the pathways identified. K.M. used the scripts developed by B.Q., Y.J., and S.S. to determine the robust clusters and to identify the reliable metabolites from this data. She also completed a pathway analysis with this information. B.Q. performed an analysis of the medications, clinical traits and age of the patient cohort to determine their involvement. K.M. drafted the manuscript (including creating the tables and supplemental materials). K.T. and R.J. helped with the coordination of patient recruitment. All coauthors provided constructive feedback and approved the manuscript. The authors have declared that there are no conflicts of interest in relation to the subject of this study.

References

1. Corvin A, Donohoe G, Hargreaves A, Gallagher L, Gill M. The cognitive genetics of neuropsychiatric disorders. *Curr Top Behav Neurosci*. 2012;12:579–613.
2. Jablensky A. The diagnostic concept of schizophrenia: its history, evolution, and future prospects. *Dialogues Clin Neurosci*. 2010;12(3):271–287.

3. Lally J, Ajnakina O, Di Forti M, *et al.* Two distinct patterns of treatment resistance: clinical predictors of treatment resistance in first-episode schizophrenia spectrum psychoses. *Psychol Med*. 2016;46(15):3231–3240.
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association; 2013.
5. Casey BJ, Craddock N, Cuthbert BN, Hyman SE, Lee FS, Ressler KJ. DSM-5 and RDoC: progress in psychiatry research? *Nat Rev Neurosci*. 2013;14(11):810–814.
6. Davison J, O'Gorman A, Brennan L, Cotter DR. A systematic review of metabolite biomarkers of schizophrenia. *Schizophr Res*. 2018;195:32–50.
7. Li C, Wang A, Wang C, *et al.* Metabolomics in patients with psychosis: a systematic review. *Am J Med Genet B Neuropsychiatr Genet*. 2018;177(6):580–588.
8. Aure MR, Vitelli V, Jernström S, *et al.*; OSBREAC. Integrative clustering reveals a novel split in the luminal A subtype of breast cancer with impact on outcome. *Breast Cancer Res*. 2017;19(1):44.
9. Sinha A, Desiraju K, Aggarwal K, *et al.* Exhaled breath condensate metabolome clusters for endotype discovery in asthma. *J Transl Med*. 2017;15(1):262.
10. Martí-Juan G, Sanroma G, Piella G; Alzheimer's Disease Neuroimaging Initiative and the Alzheimer's Disease Metabolomics Consortium. Revealing heterogeneity of brain imaging phenotypes in Alzheimer's disease based on unsupervised clustering of blood marker profiles. *PLoS One*. 2019;14(3):e0211121.
11. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13(2):261–276.
12. Trakadis YJ, Fulginiti V, Walterfang M. Inborn errors of metabolism associated with psychosis: literature review and case-control study using exome data from 5090 adult individuals. *J Inherit Metab Dis*. 2018;41(4):613–621.
13. Bouatra S, Aziat F, Mandal R, *et al.* The human urine metabolome. *PLoS One*. 2013;8(9):e73076.
14. Boehmke B, Greenwell BM. K-means clustering. In: *Hands-On Machine Learning with R*. 1st ed. Boca Raton, FL: CRC Press; 2019:488.
15. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach Learn*. 2003;52(1/2):91–118.
16. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26(12):1572–1573.
17. Xia J, Wishart DS. Using Metaboanalyst 3.0 for comprehensive metabolomics data analysis. *Curr Protoc Bioinforma*. 2016;2016:14.10.1–14.10.91.
18. Wishart DS, Feunang YD, Guo AC, *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2018;46(D1):D1074–D1082.
19. Pegg AE. Mammalian polyamine metabolism and function. *IUBMB Life*. 2009;61(9):880–894.
20. Leppik L, Kriisa K, Koido K, *et al.* Profiling of amino acids and their derivatives biogenic amines before and after antipsychotic treatment in first-episode psychosis. *Front Psychiatry*. 2018;9:155.
21. Koido K, Innos J, Haring L, Zilmer M, Ottas A, Vasar E. Taurine and epidermal growth factor belong to the signature of first-episode psychosis. *Front Neurosci*. 2016;10:331.
22. Nucifora LG, Tanaka T, Hayes LN, *et al.* Reduction of plasma glutathione in psychosis associated with schizophrenia and bipolar disorder in translational psychiatry. *Transl Psychiatry*. 2017;7(8):e1215.
23. Rigobello MP, Toninello A, Siliprandi D, Bindoli A. Effect of spermine on mitochondrial glutathione release. *Biochem Biophys Res Commun*. 1993;194(3):1276–1281.
24. Raffa M, Atig F, Mhalla A, Kerkeni A, Mechri A. Decreased glutathione levels and impaired antioxidant enzyme activities in drug-naive first-episode schizophrenic patients. *BMC Psychiatry*. 2011;11:124.
25. Xin L, Mekle R, Fournier M, *et al.* Genetic polymorphism associated prefrontal glutathione and its coupling with brain glutamate and peripheral redox status in early psychosis. *Schizophr Bull*. 2016;42(5):1185–1196.
26. Chen G, Yang D, Yang Y, *et al.* Amino acid metabolic dysfunction revealed in the prefrontal cortex of a rat model of depression. *Behav Brain Res*. 2015;278:286–292.
27. Khalil RB. Pantothenate's possible role in schizophrenia pathogenesis. *Clin Neuropharmacol*. 2012;35(6):296.
28. Chengappa KN, Turkin SR, DeSanti S, *et al.* A preliminary, randomized, double-blind, placebo-controlled trial of L-carnosine to improve cognition in schizophrenia. *Schizophr Res*. 2012;142(1-3):145–152.
29. Clelland CL, Drouet V, Rilett KC, *et al.* Evidence that COMT genotype and proline interact on negative-symptom outcomes in schizophrenia and bipolar disorder. *Transl Psychiatry*. 2016;6(9):e891.
30. Phang JM, Hu CA, Valle D. Disorders of proline and hydroxyproline metabolism. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease. Vol. II.* 8th ed. New York, NY: McGraw-Hill; 2001:1821–1838.
31. He Y, Yu Z, Giegling I, *et al.* Schizophrenia shows a unique metabolomics signature in plasma. *Transl Psychiatry*. 2012;2:e149.
32. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*. 2017;45(D1):D353–D361.
33. Smythies J. Redox mechanisms at the glutamate synapse and their significance: a review. *Eur J Pharmacol*. 1999;370(1):1–7.
34. Javitt DC. Glutamate and schizophrenia: phencyclidine, N-methyl-D-aspartate receptors, and dopamine-glutamate interactions. *Int Rev Neurobiol*. 2007;78:69–108.
35. Middleton FA, Mirnics K, Pierri JN, Lewis DA, Levitt P. Gene expression profiling reveals alterations of specific metabolic pathways in schizophrenia. *J Neurosci*. 2002;22(7):2718–2729.
36. Pettegrew JW, Keshavan MS, Panchalingam K, *et al.* Alterations in brain high-energy phosphate and membrane phospholipid metabolism in first-episode, drug-naive schizophrenics. A pilot study of the dorsal prefrontal cortex by in vivo phosphorus 31 nuclear magnetic resonance spectroscopy. *Arch Gen Psychiatry*. 1991;48(6):563–568.
37. Stanley JA, Williamson PC, Drost DJ, *et al.* Membrane phospholipid metabolism and schizophrenia: an in vivo 31P-MR spectroscopy study. *Schizophr Res*. 1994;13(3):209–215.
38. Yacubian J, de Castro CC, Ometto M, *et al.* 31P-spectroscopy of frontal lobe in schizophrenia: alterations in phospholipid and high-energy phosphate metabolism. *Schizophr Res*. 2002;58(2-3):117–122.

39. Farooqui AA. Status and potential therapeutic importance of n–3 fatty acids in neuropsychiatric disorders. In: *Beneficial Effects of Fish Oil on Human Brain*. New York, NY: Springer US; 2009:293–332.
40. Horrobin DF. The membrane phospholipid hypothesis as a biochemical basis for the neurodevelopmental concept of schizophrenia. *Schizophr Res*. 1998;30(3):193–208.
41. Smesny S, Milleit B, Hipler UC, *et al.* Omega-3 fatty acid supplementation changes intracellular phospholipase A2 activity and membrane fatty acid profiles in individuals at ultra-high risk for psychosis. *Mol Psychiatry*. 2014;19(3):317–324.
42. Demisch L, Gerbaldo H, Gebhart P, Georgi K, Bochnik HJ. Incorporation of 14C-arachidonic acid into platelet phospholipids of untreated patients with schizophreniform or schizophrenic disorders. *Psychiatry Res*. 1987;22(4):275–282.
43. Woods SW, Walsh BC, Hawkins KA, *et al.* Glycine treatment of the risk syndrome for psychosis: report of two pilot studies. *Eur Neuropsychopharmacol*. 2013;23(8):931–940.
44. Evins AE, Fitzgerald SM, Wine L, Rosselli R, Goff DC. Placebo-controlled trial of glycine added to clozapine in schizophrenia. *Am J Psychiatry*. 2000;157(5):826–828.
45. Fusar-Poli P, Berger G. Eicosapentaenoic acid interventions in schizophrenia: meta-analysis of randomized, placebo-controlled studies. *J Clin Psychopharmacol*. 2012;32(2):179–185.
46. Berger GE, Wood SJ, Wellard RM, *et al.* Ethyl-eicosapentaenoic acid in first-episode psychosis. A 1H-MRS study. *Neuropsychopharmacology*. 2008;33(10):2467–2473.
47. Bentsen H, Osnes K, Refsum H, Solberg DK, Bøhmer T. A randomized placebo-controlled trial of an omega-3 fatty acid and vitamins E+C in schizophrenia. *Transl Psychiatry*. 2013;3:e335.
48. Berger GE, Proffitt TM, McConchie M, *et al.* Ethyl-eicosapentaenoic acid in first-episode psychosis: a randomized, placebo-controlled trial. *J Clin Psychiatry*. 2007;68(12):1867–1875.
49. Fenton WS, Dickerson F, Boronow J, Hibbeln JR, Knable M. A placebo-controlled trial of omega-3 fatty acid (ethyl eicosapentaenoic acid) supplementation for residual symptoms and cognitive impairment in schizophrenia. *Am J Psychiatry*. 2001;158(12):2071–2074.
50. Emsley R, Niehaus DJ, Koen L, *et al.* The effects of eicosapentaenoic acid in tardive dyskinesia: a randomized, placebo-controlled trial. *Schizophr Res*. 2006;84(1):112–120.
51. Pawełczyk T, Grancow-Grabka M, Kotlicka-Antczak M, Trafalska E, Pawełczyk A. A randomized controlled study of the efficacy of six-month supplementation with concentrated fish oil rich in omega-3 polyunsaturated fatty acids in first-episode schizophrenia. *J Psychiatr Res*. 2016;73:34–44.
52. Arvindakshan M, Ghate M, Ranjekar PK, Evans DR, Mahadik SP. Supplementation with a combination of omega-3 fatty acids and antioxidants (vitamins E and C) improves the outcome of schizophrenia. *Schizophr Res*. 2003;62(3):195–204.
53. Pawełczyk T, Piątkowska-Janko E, Bogorodzki P, *et al.* Omega-3 fatty acid supplementation may prevent loss of gray matter thickness in the left parieto-occipital cortex in first episode schizophrenia: a secondary outcome analysis of the OFFER randomized controlled study. *Schizophr Res*. 2018;195:168–175.
54. Amminger GP, Mechelli A, Rice S, *et al.* Predictors of treatment response in young people at ultra-high risk for psychosis who received long-chain omega-3 fatty acids. *Transl Psychiatry*. 2015;5:e495.
55. Gupta M, Balakrishnan S, Pandhi P. Role of nitric oxide in experimental models of psychosis in rats. *Methods Find Exp Clin Pharmacol*. 2001;23(9):497–500.
56. Asevedo E, Cunha GR, Zugman A, Mansur RB, Brietzke E. N-acetylcysteine as a potentially useful medication to prevent conversion to schizophrenia in at-risk individuals. *Rev Neurosci*. 2012;23(4):353–362.
57. Rapado-Castro M, Dodd S, Bush AI, *et al.* Cognitive effects of adjunctive N-acetyl cysteine in psychosis. *Psychol Med*. 2017;47(5):866–876.
58. Retsa C, Knebel JF, Geiser E, *et al.* Treatment in early psychosis with N-acetyl-cysteine for 6 months improves low-level auditory processing: pilot study. *Schizophr Res*. 2018;191:80–86.
59. Conus P, Seidman LJ, Fournier M, *et al.* N-acetylcysteine in a double-blind randomized placebo-controlled trial: toward biomarker-guided treatment in early psychosis. *Schizophr Bull*. 2018;44(2):317–327.
60. Miodownik C, Lerner V, Vishne T, Sela BA, Levine J. High-dose vitamin B6 decreases homocysteine serum levels in patients with schizophrenia and schizoaffective disorders: a preliminary study. *Clin Neuropharmacol*. 2007;30(1):13–17.
61. Joshi M, Akhtar M, Najmi AK, Khuroo AH, Goswami D. Effect of zinc in animal models of anxiety, depression and psychosis. *Hum Exp Toxicol*. 2012;31(12):1237–1243.
62. Zimbrean PC, Schilsky ML. The spectrum of psychiatric symptoms in Wilson's disease: treatment and prognostic considerations. *Am J Psychiatry*. 2015;172(11):1068–1072.
63. Mortazavi M, Farzin D, Zarhghami M, Hosseini SH, Mansoori P, Nateghi G. Efficacy of zinc sulfate as an add-on therapy to risperidone versus risperidone alone in patients with schizophrenia: a double-blind randomized placebo-controlled trial. *Iran J Psychiatry Behav Sci*. 2015;9(3):e853.
64. Ding Q, Li H, Tian X, *et al.* Zinc and imipramine reverse the depression-like behavior in mice induced by chronic restraint stress. *J Affect Disord*. 2016;197:100–106.
65. Nowak G. Zinc, future mono/adjunctive therapy for depression: mechanisms of antidepressant action. *Pharmacol Rep*. 2015;67(3):659–662.
66. Richardson Andrews RC. Unification of the findings in schizophrenia by reference to the effects of gestational zinc deficiency. *Med Hypotheses*. 1990;31(2):141–153.
67. Vogeser M, Seger C. Pitfalls associated with the use of liquid chromatography-tandem mass spectrometry in the clinical laboratory. *Clin Chem*. 2010;56(8):1234–1244.