Journal of Proteomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Targeted proteomic analysis of cognitive dysfunction in remitted major depressive disorder: Opportunities of multi-omics approaches towards predictive, preventive, and personalized psychiatry

K. Oliver Schubert[a,b,*], David Stacey[a,c], Georgia Arentz[d], Scott R. Clark[a], Tracy Air[a], Peter Hoffmann[d], Bernhard T. Baune[a]

[a] Discipline of Psychiatry, Adelaide Medical School, The University of Adelaide, Adelaide, South Australia, Australia
[b] Mental Health Services, Northern Adelaide Local Health Network, Lyell McEwin Hospital, Elizabeth Vale, Australia
[c] MRC/BHF Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
[d] Biomaterials Engineering and Nanomedicine, Future Industries Institute, University of South Australia, Adelaide, Australia

ARTICLE INFO

Keywords:
Major depression
Cognitive deficits
Proteomics
Gene expression
WGCNA

ABSTRACT

In order to accelerate the understanding of pathophysiological mechanisms and clinical biomarker discovery and in psychiatry, approaches that integrate multiple –omics platforms are needed. We introduce a workflow that investigates a narrowly defined psychiatric phenotype, makes use of the potent and cost-effective discovery technology of gene expression microarrays, applies Weighted Gene *Co*-Expression Network Analysis (WGCNA) to better capture complex and polygenic traits, and finally explores gene expression findings on the proteomic level using targeted mass-spectrometry (MS) technologies. To illustrate the effectiveness of the workflow, we present a proteomic analysis of peripheral blood plasma from patient's remitted major depressive disorder (MDD) who experience ongoing cognitive deficits. We show that co-expression patterns previous detected on the transcript level could be replicated for plasma proteins, as could the module eigengene correlation with cognitive performance. Further, we demonstrate that functional analysis of multi-omics data has the potential to point to cellular mechanisms and candidate biomarkers for cognitive dysfunction in MDD, implicating cell cycle regulation by cyclin D3 (CCND3), regulation of protein processing in the endoplasmatic reticulum by Thioredoxin domain-containing protein 5 (TXND5), and modulation of inflammatory cytokines by Tripartite Motif Containing 26 (TRI26).

Significance: This paper discusses how data from multiple –omics platforms can be integrated to accelerate biomarker discovery in psychiatry. Using the phenotype of cognitive impairment in remitted major depressive disorder (MDD) as an example, we show that the application of a systems biology approach - weighted gene co-expression network analysis (WGCNA) - in the discovery phase, and targeted proteomic follow-up of results, provides a structured avenue towards uncovering novel candidate markers and pathways for personalized clinical psychiatry.

1. Introduction

Psychiatric disorders such as Major Depressive Disorder (MDD), Bipolar Affective Disorder (BPAD) and schizophrenia (SCZ), are amongst the largest contributors to all-cause disease burden in developed countries [1]. For many sufferers, these illnesses take a chronic course and can be associated with functional decline and loss of independence.

Despite decades of scientific inquiry, the molecular mechanisms driving the aetiopathogenesis of psychiatric disorders remain poorly understood, and treatments remain largely unspecific and symptomatic rather than curative (for example, see [2]). Omics techniques (e.g. genomics, transcriptomics, proteomics, metabolomics, lipidomics) represent avenues towards better understanding of processes underpinning psychiatric disorders. These techniques can simultaneously quantify large numbers of molecules in a given biological substrate, thereby uncovering disease- or treatment- related pathways and processes previously invisible to purely hypothesis-driven research. Consequently, multiple studies have been undertaken to profile post-mortem brain tissue, cerebrospinal fluid (CSF), and peripheral body

* Corresponding author at: Discipline of Psychiatry, Adelaide Medical School, Level 7, Adelaide Health and Medical Science Building, 57 North Terrace, Adelaide 5000, Australia.
E-mail address: oliver.schubert@adelaide.edu.au (K.O. Schubert).

https://doi.org/10.1016/j.jprot.2018.02.023
Received 28 October 2017; Received in revised form 29 January 2018; Accepted 16 February 2018

fluids such as blood serum or plasma of patients with psychiatric disorders. These investigations have identified a range of converging pathways and processes implicated in psychiatric disorders, such as oxidative stress, mitochondrial function, energy metabolism, the cytoskeleton, and the synapse [3–7]. However, findings have largely been unspecific to DSM-defined diagnoses, and have in many cases been reported across a range of disorders. Whether these overlaps reflect true mechanistic communalities relevant to neuropsychiatric brain conditions [8,9], or whether they are a consequence of "déjà vu" phenomena where similar findings for diverse disorders may be a reflection of technical shortcomings and biases rather than actual biological discoveries [10,11], remains unclear. Due to these difficulties, the development of clinical diagnostics for psychiatric disorders based on –omics research has remained relatively unproductive and controversial [12], and currently no protein biomarker exists to guide clinical decisions in psychiatry.

To progress the understanding of molecular patho-mechanisms and subsequent psychiatric biomarker development to the next level, novel approaches to discovery and follow-up are required. In this paper, we propose and demonstrate a multi-omics strategy that 1) identifies psychiatric phenotypes that are clinically and functionally relevant to patients but not necessarily reflected in current DSM nosology, 2) makes use of the potent and cost-effective discovery technology of gene expression microarrays, 3) employs a systems biology approach to data analysis in complex and polygenic traits, using weighted gene co-expression network analysis (WGCNA), 4) explores gene expression findings on the proteomic level using targeted mass-spectrometry (MS) technologies, and 5) selects molecular candidates for further follow up based on pathway- and network analyses and on overlapping findings in gene-expression and proteomic experiments.

We applied this approach to the investigation of the molecular profile of people suffering long-term cognitive impairments following an episode of major depression (MDD). Impairments in cognitive domains including attention, memory, executive function, and social perception occur in about 40% of patients with Major Depressive Disorder (MDD) following remission from acute episodes [13]. It has been proposed that cognitive impairment in MDD might represent a trait marker, trait-by-illness interaction marker, or endophenotype of MDD that is independent of the acute depressive state and is particularly active in a subgroup of MDD patients [14,15]. We considered the phenotype as clinically important, due to the significant consequences endured by this subgroup such as reductions in pre-morbid functioning and higher unemployment rates [16,17], and a lack of specific therapies targeting cognitive difficulties in MDD. To explore the potential molecular underpinnings of a putative "cognitive" MDD phenotype, we previously undertook gene-expression profiling of a well-matched pilot sample of remitted MDD patients with poorer ($n = 10$) versus better performance ($n = 9$) on cognitive testing [18].

In our workflow, gene expression microarray technology was used as the initial discovery platform. For shotgun experiments, gene expression arrays offer good 'value for money', covering a substantial proportion of the human transcriptome at reasonable cost. For example, the arrays used for our initial discovery experiment detected 7964 probes after data filtering and –normalization [18], about 35% of the estimated 20,110 protein-coding genes (PCGs) expressed in human tissues [19]. Because we assumed that a 'cognitive' MDD phenotype is highly complex, polygenic, and heterogeneous, we performed weighted gene co-expression network analysis (WGCNA), a hypothesis-free systems biology approach that identifies 'modules' of co-regulated – and therefore functionally related – genes in transcriptomic datasets [20]. After constructing a gene co-expression network from samples of all patients, we determined whether modules were correlated with poorer versus better cognitive performance. We additionally explored whether modules were also correlated with a continuous measure of cognitive performance in both groups. In our sample of remitted MDD patients, we identified a co-expression network of 16 distinct modules. Of these, only one module comprising 94 co-expressed genes and denoted as 'salmon' module, was significantly correlated with cognitive function. Functional analysis of the 'salmon' module implicated a down-regulation of transcripts promoting B lymphocyte proliferation, as well as a down-regulation of ribosomal S26 transcripts in MDD patients with cognitive difficulties [18].

For the current study, we sought to replicate these findings on the protein level, in blood plasma of the same cohort of patients with remitted MDD. We carried out a targeted proteomic analysis, searching for the 94 proteins previously implicated in the 'salmon' module of the gene-expression study. For proteins detected by Middle Band Collision-Induced Dissociation (CID) mass spectrometry (MS), we calculated a module eigengene (i.e., the first principal component of the standardized protein expression profile of a given module), and tested its correlation with patients' cognitive performance. Further, we explored the relationship between transcriptomic and proteomic data using QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity).

2. Materials and methods

2.1. Study participants

Patients with MDD and healthy controls were recruited for the Adelaide Cognitive Function and Mood Study (CoFaMS, HREC RAH protocol No 111230). Nineteen patients with remitted MDD, were matched for race, gender, age, Hamilton Depression Scale score, presence of psychotic features during any depressive episode, educational status, annual household income, current drug- and alcohol use, and current use of medication, as described previously [18] (Table 1). Remitted MDD status was assessed with the MINI-6.0.0 Neuropsychiatric Diagnostic Interview [21]. To allow for binary analysis, a group of 'poorer' cognitive functioning (MDD-Cog-, $n = 10$), and a group of 'better' cognitive functioning (MDD-Cog+, $n = 9$) was formed based on total scores on the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) [22]. For the Cog- group, we selected patients with RBANS scores below the 40th percentile, for the Cog+ group patients with RBANS scores above the 55th percentile. To achieve sufficient contrast between the groups, we excluded participants with scores around the 50th percentile.

Table 1
Overview of demographic characteristics of the study groups for Middle-band CID MS, including potential confounders for neurocognitive performance.
(reproduced with permission from Schubert et al., 2016 [18]).

	MDD-Cog$^-$ cognitive deficits (< 40th centile RBANS)	MDD-Cog$^+$ high-normal cognitive functioning (> 55th centile RBANS)
Number	10	9
Race	9 Causcasian, 1 Asian	8 Caucasian, 1 Asian
Gender	4M, 6F	4M, 5F
Mean age and range	35.1 (18–71)[a]	43.4 (18–69)[a]
Psychotic features	0	0
Years of education	12.8	15.1**
Enrolled tertiary education	6	7
Year 12 completed	10	9
Annual household income (less than $40 kpa)	3	6
HAM-D score (mean ± SD)	9.1 ± 5.7	11.1 ± 7.3[a]
Alcohol abuse (current)	0	0
Drug abuse (current)	0	0
Using antidepressants	3	4
Using antipsychotics	0	0
Using mood stablisers	0	0

[a] No significant difference between groups as per *t*-test.
** Significant at $p < 0.05$.