Online Research @ Cardiff

information services
gwasanaethau gwybodaeth

Blood-Based Protein Changes in Childhood Are Associated With Increased Risk for Later Psychotic Disorder: Evidence From a Nested Case–Control Study of the ALSPAC Longitudinal Birth Cohort

Jane A. English[1], Lorna M. Lopez[1], Aoife O'Gorman[1], Melanie Föcking[1], Magdalena Hryniewiecka[1], Caitriona Scaife[2], Sophie Sabherwal[1], Kieran Wynne[2], Patrick Dicker[3], Bart P. F. Rutten[4], Glynn Lewis[5], Stanley Zammit[6,7], Mary Cannon[1], Gerard Cagney[2], and David R. Cotter*,[1]

[1]Department of Psychiatry, Royal College of Surgeons in Ireland, Dublin, Ireland; [2]School of Biomolecular and Biomedical Science, Conway Institute, University College Dublin (UCD), Dublin, Ireland; [3]Department of Epidemiology and Public Health, Royal College of Surgeons in Ireland, Dublin, Ireland; [4]Department of Psychiatry & Neuropsychology, School for Mental Health and Neuroscience, Division of Neuroscience, Maastricht University, Maastricht, The Netherlands; [5]Division of Psychiatry, UCL, London, UK; [6]MRC Centre for Neuropsychiatric Genetics & Genomics, Cardiff University, Cardiff, UK; [7]School of Social and Community Medicine, University of Bristol, Bristol, UK

*To whom correspondence should be addressed; RCSI Education & Research Centre, Beaumont Hospital, Beaumont, Dublin 9, Ireland; tel: 01-809-3862, fax: 01-809-3741, e-mail: drcotter@rcsi.ie

The identification of early biological changes associated with the psychotic disorder (PD) is important as it may provide clues to the underlying pathophysiological mechanisms. We undertook the first proteomic profiling of blood plasma samples of children who later develop a PD. Participants were recruited from the UK Avon Longitudinal Study of Parents and Children (ALSPAC) cohort who also participated in psychiatric assessment interviews at age 18. Protein expression levels at age 11 were compared between individuals who developed PD at age 18 ($n = 37$) with population-based age-matched controls ($n = 38$). Sixty out of 181 plasma proteins profiled were found to be differentially expressed ($P < .05$) in children with an outcome of the PD. Thirty-four of these proteins were found to be differentially expressed following correction for multiple comparisons. Pathway analysis implicated the complement and coagulation cascade. A second, targeted proteomic approach was used to verify these findings in age 11 plasma from subjects who reported psychotic experiences at age 18 ($n = 40$) in comparison to age-matched controls ($n = 66$). Our findings indicate that the complement and coagulation system is dysregulated in the blood during childhood before the development of the PD.

Key words: ALSPAC/complement/coagulation/plasma/proteomics/psychotic disorder/schizophrenia

Introduction

There is firm evidence that the early identification and treatment of subjects with psychotic disorder (PD)

significantly improve their clinical outcome.[1] As clinical characteristics alone are of limited predictive value, there is a now a pressing need for further enhancement of predictive models to allow the earliest identification of subjects who may later develop psychosis.[2,3]

Blood-based precursors of the PD may offer the opportunity to enhance predictive models and may provide important clues for our understanding of pathophysiological cascades which may later be expressed as psychosis or PD. Our recent review of blood biomarkers in the PD schizophrenia points to a convergence of pathophysiological mechanisms that involve the acute-phase response, glucocorticoid receptor signalling, coagulation, and lipid and glucose metabolism.[4] Furthermore, inflammatory cytokines, chemokines, and growth factors have been assessed in the blood during the perinatal periods and during childhood in subjects who subsequently developed schizophrenia, and in those with a first-episode psychosis.[5–9] These studies have, in general, demonstrated a picture of enhanced inflammatory tone during and preceding psychosis. While the basis of these changes is not clear, numerous risk factors for schizophrenia such as genetic background but also exposures to abuse, maternal stress during pregnancy, prenatal famine, obstetric complications, and adolescent cannabis use have all been described and hypothesized to lead to raised inflammatory tone.[10] Post-mortem brain studies, albeit possibly confounded by many variables, have furthermore underpinned the evidence for a role of inflammation, suggesting that this process is involved during early but also later stages of the disorder.[11–13]

One of the main challenges of the field is to identify differential profiles which can be measured prior to the clinical onset of the PD. Previous studies of the ALSPAC cohort, a prospective general population cohort based in Bristol area in South West England have characterized subgroups of subjects who developed PD and Psychotic Experiences (PEs)[14] at age 18. These groups showed alterations in cortical white matter microstructure,[15] working memory,[16] and raised inflammatory markers in childhood[5] in subjects later developing PD and PE. Supported by this evidence, we employed a discovery proteomic study on blood plasma samples at age 11 in order to yield insights into the underlying protein pathways dysregulated in children with PD outcomes at age 18. As expression of PEs can be considered as a liability phenotype for PDs and may involve the same etiopathogenesis,[17] we next quantified candidate proteins in blood samples of subjects at age 11 in a separate subsample from the ALSPAC cohort, and compared subjects who reported PEs at age 18 with matched control subjects without PEs at age 18 as an verification study.

Methods

For extended materials and methods, please refer to online supplementary methods.

Participants

The ALSPAC cohort is a prospective general population cohort, and a rich resource of demographic, environmental, and clinical data on the individuals involved.[18] Written informed consent was obtained prior to taking the plasma samples. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Please note that the study website contains details of all the data that is available through a fully searchable data dictionary (http://www.bristol.ac.uk/alspac/researchers/access).

Measures of PE and PD

PEs were identified at 11 and 18 years through the face-to-face, semi-structured Psychosis-Like Symptom (PLIKS) interview[14] conducted by trained psychology graduates in assessment clinics, and were coded according to the definitions and rating rules for the Schedules for Clinical Assessment in Neuropsychiatry, Version 2.0 (Organisation 1994). Interviewers rated PEs as not present, suspected or definitely psychotic. Cases of PD were defined as individuals with definite PEs that were not attributable to the effects of sleep or fever and when the PE occurred at least once per month over the past 6 months and caused severe distress, had a very negative effect on social/occupational function, or led to help seeking from a professional source.[14] All subjects categorized at the interview as having a PD met diagnostic criteria for PDs as defined in both DSM-IV and ICD-10, given that they had regular psychotic phenomena that were causing them severe distress or substantially impaired functioning.[14] Control samples from age-matched individuals without suspected or definite PEs, or PD at age 18 were randomly selected to match cases (table 1). With regard to psychotropic drug use, 8 of the 37 subjects (21.6%) with PD and who provided plasma samples were on psychotropic medication at age 18. No subjects reported psychotropic drug use at age 11. Subjects categorized as having disorder in our study only included those whose PEs were not attributed to drug use.[14]

Study Design

We undertook a nested case-control study of the ALSPC cohort and chose to assess all available plasma samples from age 11 children with outcomes of PD or PE at age 18. Available plasma samples from controls of age-matched individuals were then randomly selected to arrive at the sample size for each analysis. The present study consisted of analyses of two datasets harboring mass-spectrometry-based proteomic profiles. (1) Discovery proteomic

Table 1. Descriptive Information for ALSPAC Subjects Included in the PD and PE Studies

	Psychotic Disorder (PD) Discovery Analysis		Psychotic Experiences (PE) Verification Analysis	
	Cases	*Controls*	*Cases*	*Controls*
Number	37	38	40	66
Gender	29F, 8M	28F, 10M	22F, 18M	27F, 39M
BMI at age 11, Mean (SD)	18.24 (3.39)	18.066 (2.69)	18.12 (2.72)	17.75 (2.52)
Ethnicity	28 W, 3 NW, 6 NA	37 W, 1 NA	37 W, 1 NW, 2 NA	63 W, 0 NW, 3 NA
PLIKS at age 12	20 none, 8 suspected, 9 definite	38 none	40 none	60 none, 6 NA
PLIKS at age 18	37 definite clinical	None	29 = definite nonclinical, 11 = sleep/fever	66 none

Note: For gender: F, female; M, male. Body mass index (BMI) at age 12 is reported, where missing BMI variables were replaced with the mean according to gender. For ethnicity: W, white; NW, non-white; NA, missing. PLIKS at age 12 and age 18 are reported, however, in this analyses, we used PLIKS at age 18 as the main outcome measure for our proteomic analysis.

analysis was performed on $n = 37$ individuals with an outcome of PD at age 18 vs $n = 38$ age-matched controls without PD or PE (table 1). (2) Hypothesis-driven proteomic analysis to confirm the discovery protein findings was performed on $n = 40$ individuals with an outcome of PE at age 18 vs $n = 67$ controls without PD or PE (table 1).

Blood Collection

For all participants, blood samples from nonfasting individuals were collected at approximately 11 years of age. Blood was collected in 7.5 ml Plasma Lithium-Heparin S-Monovette tubes (Sarsted). Once collected, samples were stored on ice for a maximum of 90 min until processed. After centrifugation, the plasma was stored in aliquots at −80C. All samples underwent a single freeze-thaw cycle to allow aliquotting prior to the study. The standard quality of the plasma samples was ensured by assessing the overall MS protein profile to facilitate the identification of outlier protein expression profiles (see supplementary figure S1).

High-Abundance Protein Depletion of Plasma Samples

To improve the dynamic range for proteomic analysis, 40 µl of plasma from each case was immunodepleted of the 14 most abundant proteins (alpha-1-antitrypsin, A1-acid glycoprotein, serum albumin, alpha2-macro-globulin, apolipoprotein A-I, apolipoptrotein A-II, complement C3, fibrinogen alpha/beta/gamma, haptoglobin, IgG A, IgG G, IgG M, transthyretin, and serotransferrin) using the Agilent Hu14 Affinity Removal System (MARS) coupled to a high-performance liquid chromatography (HPLC) system[19] (supplementary methods).

Sample Preparation for Mass Spectrometry

Protein digestion and peptide purification was performed as previously described,[20] and are further detailed in supplementary methods.

Discovery Proteomic Analysis in PD

For the discovery proteomic analysis in PD (37 cases vs 38 controls; see table 1), 5 µl from each sample was injected on a Thermo Scientific Q-Exactive mass spectrometer, connected to a Dionex Ultimate 3000 (RSLCnano) chromatography system, and operated in data dependent analysis (DDA) mode, for label-free LC-MS/MS, which is standard proteomic profiling method for case/control clinical investigations.[20–22] Please refer supplementary methods for further details.

Hypothesis-driven Proteomic Analysis in PE

For hypothesis-driven proteomic analysis of candidate proteins in a second subset of participants, we used targeted-data independent acquisition (DIA) analysis. Targeted-DIA is currently the gold standard for high-throughput quantification of protein targets.[23–25] For targeted-DIA in the PE subset of participants (40 cases vs 66 controls; table 1), 5 µl of each sample was injected on the Thermo Scientific Q-Exactive, and data was acquired in DIA mode. The DIA isolation scheme and multiplexing strategy were based on that from Egertson et al., in which five separate 4-m/z isolation windows are analyzed per spectrum[26,27] (supplementary methods).

Bioinformatics and Statistical Analysis

Discovery Proteomic Analysis in PD. Label-free quantification for our discovery DDA files generated from the PD subjects was performed in Max Quant (v1.5.2.8),[28,29] as described.[20] Only proteins present in >80% of samples in at least one group were taken forward for quantification (supplementary table 1; $n = 181$ proteins), whereby the data were log2 transformed and normalized (supplementary figure S1).[30] Data management and analyses were then performed in R version 3.0.1 (2013-05-16). Missing proteomic data was replaced with a random value from the normal distribution of each protein that was created separately for cases and controls.[31] Linear regression analyses were performed on all proteins with the subject sample as a random effect, adjusting for gender and BMI at age 11. We adjusted for BMI as the prevalence of obesity in people with mental illness has been reported to higher than the general population,[32,33] and recent evidence suggests childhood BMI between the ages 7 and 13 years is associated with risk of schizophrenia.[34] A 5% false discovery rate (FDR) threshold was applied using the "fdrtool" package in R, as previously described.[35,36] The demographic and clinical data were tested for differences between case and control group using the Fisher's Exact test.

Hypothesis-Driven Proteomic Analysis in PE. All DIA files from the PE analysis (see table 1 for groups) were analyzed in Skyline (V3.5.0; https://skyline.gs.washington.edu), as detailed by Egertson et al.[26,27] We aimed to verify; (1) the specific proteins identified as differentially expressed following FDR from our "***Discovery proteomic analysis in PD"*** and (2) the broader set and proteins from the complement and coagulation pathways which were generally implicated from the *Discovery proteomic analysis in PD'*. This was done in a separate ALSPAC subgroup of participants with a milder phenotype of PE ($n = 40$) but not PD at age 18, compared to controls ($n = 66$; table 1). For a full list of the fragments targeted and quantified please refer to supplementary table S2. All peptides and associated fragment ions were visually checked, and peak editing was undertaken where necessary (supplementary document 1). Preprocessing and statistical analysis of the fragment level data were

undertaken in mapDIA[37] (supplementary table S3), as detailed in supplementary methods.

Pathway Analysis

KEGG pathway analysis is traditionally performed on all proteins that are differentially expressed ($P < .05$) (http://www.genome.jp/kegg/pathway.html) in order to identify coordinated changes in pathways and cellular processes at the level of the proteome, and this was therefore undertaken on all proteins differentially expressed in the arising from the PD study. The STRING database (http://string-db.org/)[38] was used to analyze the protein–protein interactions and pathway relationships within our dataset.

Results

Discovery Proteomic Analysis in PD

For the discovery proteomic analysis in PD (37 cases vs 38 controls; table 1), we identified 345 proteins (FDR < 0.01), of which 181 were present in >80% of samples and were taken forward for analysis (supplementary table 1). Sixty proteins out of 181 were significantly differentially expressed between PD cases and controls ($P < .05$), following adjustment for gender and BMI as covariates (see results, table 2). Thirty-four proteins (A2M, TNXB, IGHM, CLEC3B, BCHE, GSN, ECM1, IGF2, C4BPB, CTBS, SERPINA7, VCAM1, LUM, GPLD1, HABP2, IGHG3, SEPP1, CFH, SERPINF1, F5, F9, IGFBP3, F12, CFD, CFI, LAMP2, CRTAC1, PCOLCE, AFM, PLG, COL6A3, C7, APOA2, CD109) remained significant after FDR (table 2). The full list of proteins profiled, and those that were differentially expressed between groups is provided in supplementary table 1.

Pathway Analysis

To identify potential protein pathways and processes whose expression is co-ordinately altered in PD, we uploaded the 60 significant proteins to KEGG for pathway analysis. The top pathway implicated was the "Complement and Coagulation Cascade" (figure 1a; Fisher's exact test $P = 8.56 \times 10^{-32}$), followed by *Staphylococcus aureus* infection and *Pertussis* infection protein pathways (figure 1a). In addition, when the same proteins were analyzed using the STRING algorithm, the set of coagulation and complement proteins were found to cluster together as subnetworks (figure 1b), providing further evidence of coordinated perturbation of these pathways.

Hypothesis-Driven Proteomic Analysis in PE

These analyses aimed to verify and extend the findings of the discovery analysis of PD in a separate ALSPAC subgroup of participants with a milder phenotype of PE ($n = 40$) but not PD at age 18, compared to controls ($n = 66$; table 1). Two PE cases were excluded from the bioinformatics analysis due to poor chromatographic profiles.

We used a targeted-DIA approach (see supplementary document 1) to quantify the levels of the 34 FDR significant proteins in the PE group. Twenty-seven of the 34 FDR candidates had peptides suitable for targeted-DIA analysis (supplementary table S3). We also used a targeted-DIA approach to target the proteins which mapped to the complement and coagulation cascade in PD but which were not in the list of the 34 FDR significant proteins (table 2). Thus of the total of 19 proteins which mapped to the complement and coagulation cascade (figure 1), we had already assessed 10 when undertaking our targeted-DIA assessment of the 34 FDR significant proteins, leaving 9 proteins (significant at $P < .05$) which mapped to this process (figure 1), and we assessed these 9. Therefore a total of 36 protein candidates (27 + 9) were quantified in the hypothesis-driven PE study. Out of the 36 targets, we confirmed the differential expression of 10 proteins using a Bayesian latent variable model at a 5% FDR threshold (C4BPA, A2M, PLG, AFM, CFH, IGHG3, VTN, GSN, C4BPB, and IGHM; supplementary table S3). The protein level DIA data illustrates that the direction of protein expression changes in the PD and PE phenotypes compared to controls was the same for all of these 10 proteins (figure 2).

Discussion

An urgent challenge in the field of psychosis is to understand the pathophysiological basis of psychosis so that early prediction can be realized and earlier and more effective prevention and treatment options can be considered. We have undertaken the first discovery proteomic profiling study in early adolescence in order to identify differential plasma protein expression preceding PD. Using a unique prospective cohort, we investigated blood plasma samples obtained from children at age 11 who went on to develop PD at age 18. In the samples collected at age 11, we identified a set of 34 proteins significantly differentially expressed between children who developed PD vs age-matched population-based controls (table 2). Seven of these 34 proteins (A2M, APOA2, C4BPB, CLECB3, GPLD1, IGHM, IGFBP3) have previously been implicated in blood-based biomarker studies of patients diagnosed with the PD schizophrenia.[4] These proteins, therefore, might have potential value in predicting the risk of and in contributing to future pathophysiological models for PD. However, this will require robust identification of markers, calibration of risk prediction models, and testing in independent cohorts.

The information that can be derived from the broader pathway- or network-based picture of differential protein expression (rather than the individual proteins) is relevant to understanding the biological basis of the PD. Pathway

Table 2. Differential Protein Expression in Psychotic Disorder

Gene Names	Protein Names	Protein ID	P Value	FC in PD	FDR
A2M	Alpha-2-macroglobulin	P01023	2.35E-05	−1.36	0.002452739084173
TNXB	Tenascin-X	P22105	6.78E-05	1.60	0.003010022721392
IGHM	Ig mu chain C region	P01871	8.93E-05	−1.49	0.003100233074614
CLEC3B	Tetranectin	P05452	0.000238536918096	1.21	0.0055271499724 27
BCHE	Cholinesterase	P06276	0.00028867379359	1.22	0.006016201946107
GSN	Gelsolin	P06396	0.00118433 1262904	1.13	0.013741065106703
ECM1	Extracellular matrix protein 1	Q16610	0.00156813706243	1.23	0.015289744574204
IGF2	Insulin-like growth factor II	P01344	0.00185038627232	1.26	0.016146280605035
C4BPB	C4b-binding protein beta chain	P20851	0.001968111624603	−1.31	0.016452583474952
CTBS	Di-N-acetylchitobiase	Q01459	0.002116574187167	1.23	0.016804037125592
SERPINA7	Thyroxine-binding globulin	P05543	0.002123057983554	1.17	0.016818582457348
VCAM1	Vascular cell adhesion protein 1	P19320	0.002249659902707	1.34	0.01709034422445
LUM	Lumican	P51884	0.002278966387226	1.19	0.017150104571853
GPLD1	Phosphatidylinositol-glycan-specific phospholipase D	P80108	0.00234614616507	1.42	0.017282940815611
HABP2	Hyaluronan-binding protein 2	Q14520	0.002538350481953	1.17	0.017633778389896
IGHG3	Ig gamma-3 chain C region	P01860	0.002874794375478	−1.27	0.018546283526309
SEPP1	Selenoprotein P	P49908	0.003146451379488	1.21	0.019193243065463
CFH	Complement factor H	P08603	0.003414262231026	1.13	0.019765573474774
SERPINF1	Pigment epithelium-derived factor	P36955	0.003757243579115	1.13	0.02047865413338
F5	Coagulation factor V	Q28107	0.004067608434137	1.27	0.021055629451341
F9	Coagulation factor IX	P00740	0.004334224194447	1.13	0.021506865940923
IGFBP3	Insulin-like growth factor-binding protein 3	P17936	0.007644758410206	1.17	0.032388279082007
F12	Coagulation factor XII	P00748	0.007824122898499	1.19	0.032887687197417
CFD	Complement factor D	P00746	0.008444673418423	1.35	0.034556546176356
CFI	Complement factor I	P05156	0.010003939954609	1.12	0.038384336560269
LAMP2	Lysosome-associated membrane glycoprotein 2	P13473	0.011818700619699	1.19	0.042278849036246
CRTAC1	Cartilage acidic protein 1	Q9NQ79	0.012282687547432	1.23	0.043191407027815
PCOLCE	Procollagen C-endopeptidase enhancer 1	Q15113	0.012858286614664	1.19	0.044281235146809
AFM	Afamin	P43652	0.014449995361757	1.12	0.047073156606685
PLG	Plasminogen	P00747	0.014808887658904	1.14	0.047661474616677
COL6A3	Collagen alpha-3(VI) chain	P12111	0.015123384190082	1.35	0.048165491709375
C7	Complement component C7	P10643	0.015253718180595	1.14	0.048371295829491
APOA2	Apolipoprotein A-II	P02652	0.015768184543684	−1.35	0.049166640144774
CD109	CD109 antigen	Q6YHK3	0.016302896370883	1.23	0.049965576376137
ACTG1	Actin, cytoplasmic 2	P63261	0.01842844064171	1.23	0.052893098007721
CD14	Monocyte differentiation antigen CD14	P08571	0.018648336393252	1.14	0.05317487598941
KLKB1	Plasma kallikrein	P03952	0.019913520723869	1.11	0.054727223568396
KNG1	Kininogen-1	P01042	0.021034485331523	1.11	0.056011643691759
F5	Coagulation factor V	P12259	0.021375886441285	1.14	0.056387118262621
SERPINF2	Alpha-2-antiplasmin	P08697	0.024325248700793	1.09	0.059359684835652
C1S	Complement C1s subcomponent	P09871	0.024366317843667	1.12	0.059397936011767
F2	Prothrombin	P00734	0.024740545362472	1.12	0.059742847153768
APOA4	Apolipoprotein A-IV	P06727	0.025696265563249	1.13	0.060594940495585
IGKV2D-40	Ig kappa chain V-II region TEW	A0A075B6R1	0.025902708420233	−1.20	0.060773774727794
LYVE1	Lymphatic vessel endothelial hyaluronic acid receptor 1	Q9Y5Y7	0.026443478521856	1.24	0.061233787247234
F13A1	Coagulation factor XIII A chain	P00488	0.027407334083503	1.14	0.062086047473999
S100A8	Protein S100-A8	P05109	0.030888403899574	1.24	0.065695245585865
ANPEP	Aminopeptidase N	P15144	0.030918599644144	1.22	0.065724623760281
VTN	Vitronectin	P04004	0.031818887929209	1.10	0.066586456714209

Table 2. *Continued*

Gene Names	Protein Names	Protein ID	*P* Value	FC in PD	FDR
C4BPA	C4b-binding protein alpha chain	P04003	0.034467677540817	−1.17	0.068973446582385
APOH	Beta-2-glycoprotein 1	P02749	0.035328247276582	1.13	0.06970463250612
C6	Complement component C6	P13671	0.036921822018089	1.10	0.071005866471262
C1R	Complement C1r subcomponent	P00736	0.03697367472293	1.13	0.07104710171526
CLU	Clusterin	P10909	0.037047254149251	1.09	0.071105498589918
COMP	Cartilage oligomeric matrix protein	P49747	0.038053606817258	1.27	0.071890746567978
CPN2	Carboxypeptidase N subunit 2	P22792	0.039032541881775	1.11	0.07263130677403
ADIPOQ	Adiponectin	Q15848	0.041918375605006	1.28	0.076309890815356
ATRN	Attractin	O75882	0.045675287259397	1.09	0.080866376756604
VASN	Vasorin	Q6EMK4	0.047805918283872	1.15	0.083340993340051
RARRES2	Retinoic acid receptor responder protein 2	Q99969	0.048852426112363	1.15	0.084528878416831

Note: Label-free discovery proteomic was used to compare age 11 plasma samples from children with an outcome of PD at age 18 with age-matched healthy controls (n = 75). Sixty proteins were identified as significantly differentially expressed between cases and controls (ANCOVA, $P < .05$), following adjustment for gender and BMI as covariates (supplementary table 1). The gene name, protein name, accession, ANCOVA adjusted P value, fold change (FC) in disorder, and false discovery rate (FDR) values are listed for the 60 significant proteins which are sorted by FDR.

analysis identified the Complement and Coagulation Cascade as the most significant process altered (figure 1), with a total of 19 proteins involved in the process being differentially expressed between groups. These complement and coagulation proteins included 3 decreased proteins (A2M, C4BPA, C4BPB) and 16 increased proteins (C1R, C1S, C6, C7, CFD, CFH, CFI, F12, F13A1, F2, F5, F9, KLKB1, KNG1, PLG, SERPINF2) in children who developed PD. Due to significant overlap in proteins that mapped to the other top infection related processes in *Staphylococcus* and *Pertussis* infection, we focused our attention on complement and coagulation cascade. We followed the above discovery study with a hypothesis-driven proteomic study of children who went on to develop PEs (rather than PD) at age 18. In this group, we confirmed the differential expression of 10 proteins which were altered in the PD group, and all of these changes were in the same direction as in the PD group (figure 2). Six of these targeted proteins were implicated in complement and coagulation (A2M, C4BPA, C4BPB, CFH, PLG, VTN), two in the immune system (IGHM and IGHG3), and one each in Vitamin E binding (AFM), and synaptic elimination and myelination (GSN). As PEs is a milder phenotype that PD, with about 8% of individuals ultimately transitioning to PD,[13] we consider these results to be supportive of our main findings.

The complement and coagulation cascade is part of the innate immune defense against infection, and the role of these processes in the pathophysiology of schizophrenia has gained attention in recent years. After decades of intensive research, collaborative GWAS efforts identified the strongest risk association for schizophrenia in the major histocompatibility complex (MHC)—best known

for its role in immunity.[39] Sekar et al[13] have recently shown that within this MHC region the strongest risks are associated with the complement 4 (C4) gene, which reflects the heritable impairment in the complement system in schizophrenia. In this study, we found 9 complement proteins dysregulated in children preceding the onset of PD, including decreased C4BPA and C4BPB, and increased C1R, CFD, C1S, CFI, C6, C7, and CFH. The C4BPA and C4BPB proteins assemble together forming C4b-binding protein,which regulates the activation of the complement cascade, and also has a regulator role in the coagulation system. The coagulation system was also largely increased in expression in children preceding the onset of PD, including increased KNG1, PLG, KLKB1, F12, F9, F5, F13A1, SERPINF2, F2, and decreased A2M. In support, Li et al[40] also reported the complement and coagulation cascade as the most significant pathway implicated in plasma samples of drug naïve schizophrenia patients, and again reported on the reduced expression of C4BPB. Other blood-based proteomic studies of first-episode psychosis support these findings.[19,41,42] Dysfunction in coagulation has also been reported previously in schizophrenia[4,43,44] and more recently by Hoirisch-Clapauch et al,[45] who observed that schizophrenia patients on chronic warfarin therapy showed remission of psychotic symptoms. The complement system is linked functionally with inflammation,[46] and within the brain has important functions in the regulation of synaptic plasticity.[13,47–49] Our findings are thus in keeping with the literature implicating inflammation in major psychiatric disorders[10,50] and with potential mechanisms involved in complement and altered synaptic plasticity.[13,51–53] However, the functional relationship between
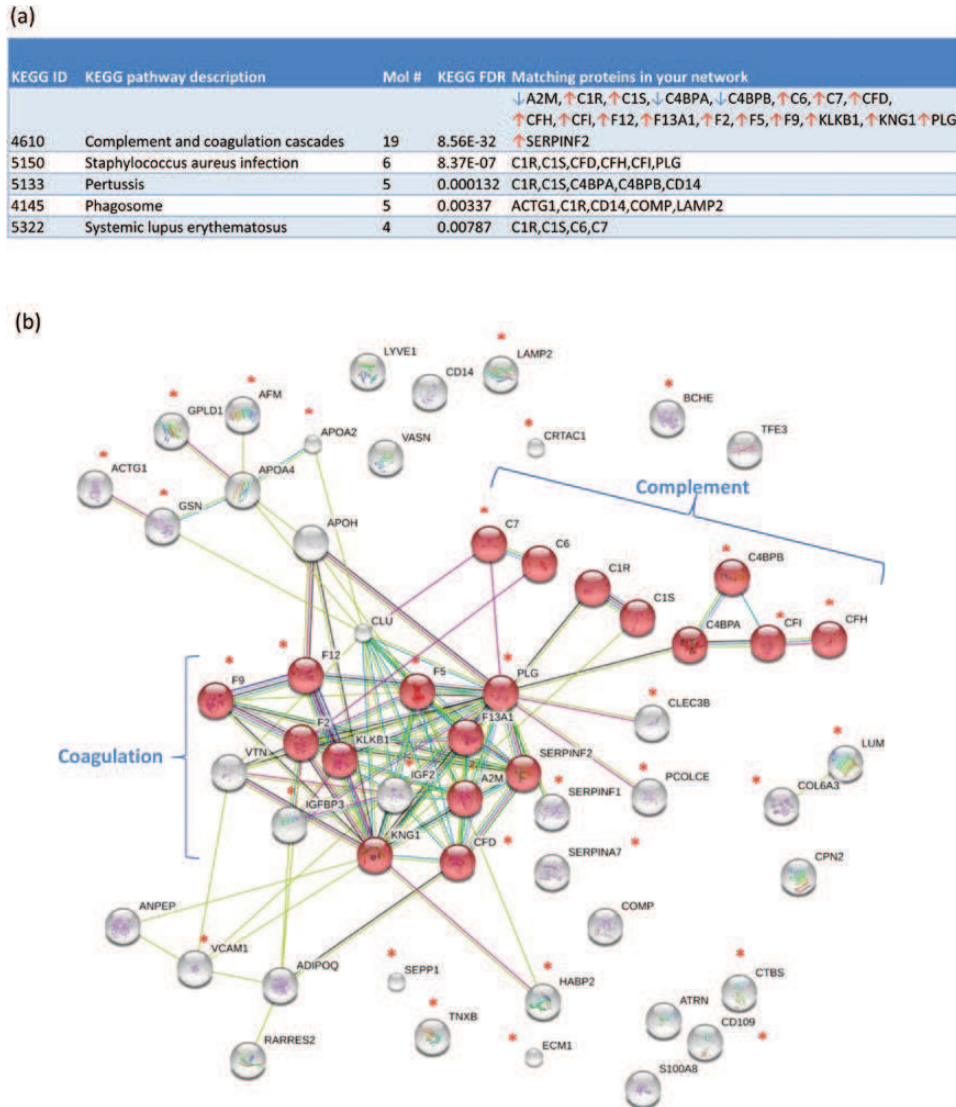
Fig. 1. Differentially expressed proteins converge on complement and the coagulation cascade. (a) KEGG pathway analysis 60 proteins as differentially expressed between cases ($n = 37$) and controls ($n = 38$) identified the complement and coagulation cascade as the top process implicated (19 proteins; false discovery rate (FDR) 8.56e-32; (b) STRING network analysis provided strong evidence for co-regulation among coagulation and complement proteins, with the average interactions per protein being 3.65 (ie, *average node degree*) for 57 of the 60 proteins mapped. Proteins that were significant at the FDR level are marked with asterisks.

complement activation within the blood and the complement activation within the brain has not been elucidated, and hence we cannot draw a direct inference between our findings and brain function. In the light of our findings, future studies need to address this issue.

Our study has several limitations which should be mentioned. Firstly, it was based on a unique characterized cohort, and we could not access an exactly similar age-matched sample in which we could perform a direct replication of the discovery analyses. However, we obtained supportive complementary data for our hypothesis-driven study using age 11 samples from a subject with PE outcomes, rather than PD outcomes, from the same cohort. Nevertheless, it has been argued that PEs form a measurable liability phenotype for PD and that PEs share (at

least some of) the risk factors of PD, suggesting that their etiopathogenesis may involve similar biological cascades. Another potential weakness is that the subjects investigated in this study were age 18 at the time of last clinical assessment and therefore have not yet gone through the full period of risk for developing a PD and hence our sample is not representative of all individuals with the PD. While future work will be required to address this, it is notable that our findings find significant support from the plasma proteomic biomarker literature of schizophrenia which has implicated the complement and coagulation pathways very consistently.[4,17,29,30,31] Finally, while various approaches for bioinformatics analyses of DIA datasets are still under development, the methods we used are in line with recent advances in the bioinformatic
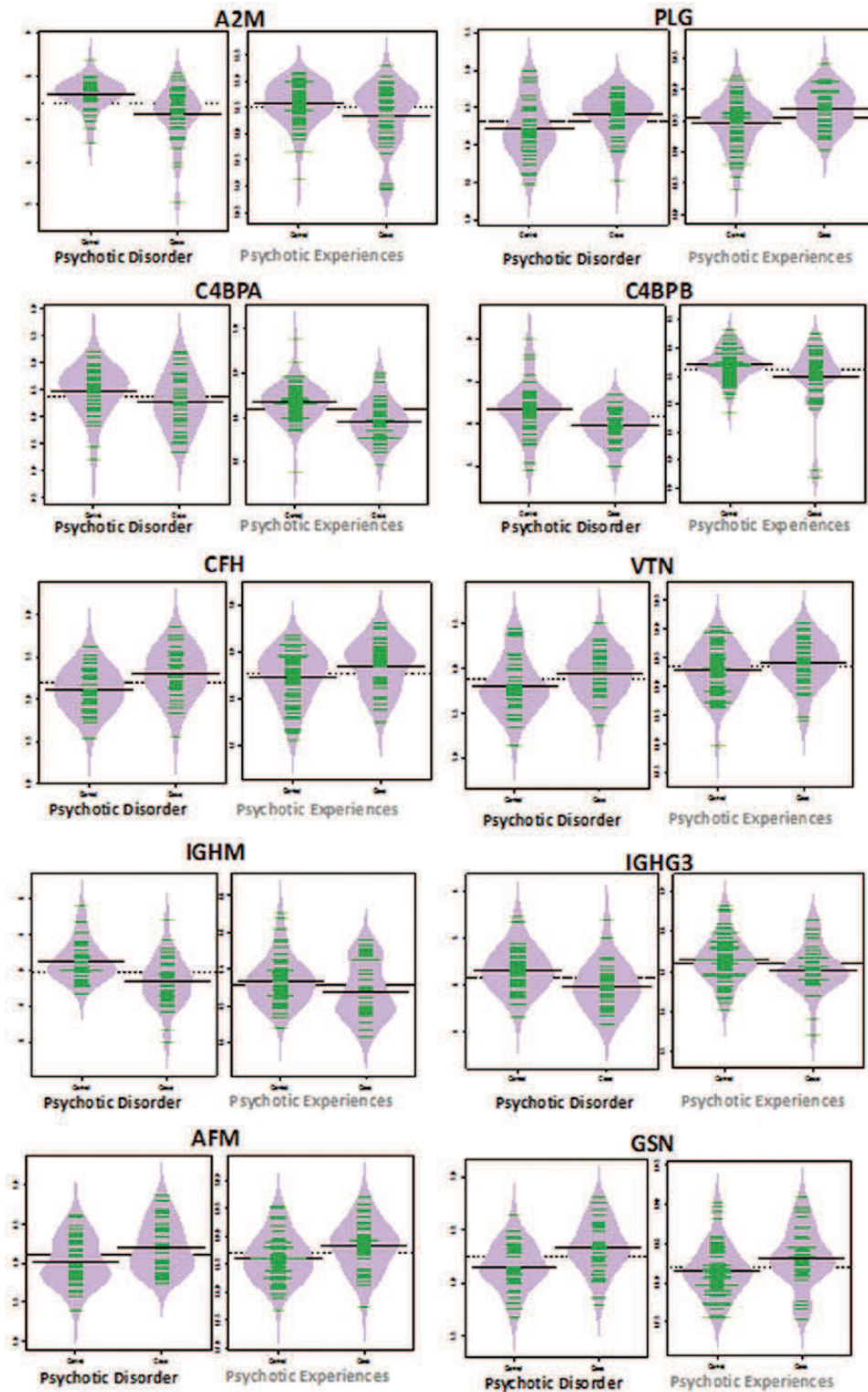
Fig. 2. Bean plots of proteins identified as altered in the same direction in both the psychotic disorder (PD) and psychotic experience (PE) cohorts. Bean plots represent the log-transformed intensities in protein expression (*y*-axis) between cases and controls (*x*-axis) for both the PD (*n* = 75) discovery proteomic analysis (left), and PE (*n* = 106) hypothesis-driven proteomic analysis (right). Please note the scale on the *y*-axis differs between the PD and PE cohorts because different quantification methods were used to assess protein expression in each cohort (ie, DDA vs DIA data). Nevertheless, bean-plots representing the expression changes in complement and coagulation proteins (A2M, PLG, C4BPA, C4BPB, CFH, and VTN), and others (IGHM, IGHG3, GSN, and AFM) are altered in the same direction in both cohorts. These protein candidates may represent persistent early hallmarks of psychotic outcomes at age 18 in the general population.

handling of DIA data by Teo et al.[37] This tool considers the fragment level data as "repeated measurements" of the abundance of peptides and thus applies a sophisticated hierarchical Bayesian algorithm[37] for the analysis of DIA data (supplementary table 3).

In conclusion, our study which is the first to undertake a proteomic analysis of plasma in childhood from subjects who develop PD in early adulthood has identified alterations in the complement and coagulation process. These changes are also present, but to a lesser extent, among subjects who go on to develop PE in early adulthood, suggesting that our findings may represent early changes associated with a vulnerability to psychosis. Further clarification of the roles of these proteins and processes will be obtained from future studies which will test how baseline levels of these proteins in subjects at ultra-high risk for PD relate to the varied clinical outcomes.[54] Overall, our findings are indicative of the presence of a pathophysiological process long before the development of the disorder, and suggest that blood-based profiles may be tested in larger population for their value in predicting later PD.

Supplementary Material

Supplementary data are available at *Schizophrenia Bulletin* online.

References

1. Larsen TK, Melle I, Auestad B, et al. Early detection of psychosis: positive effects on 5-year outcome. *Psychol Med*. 2011;41:1461–1469.
2. Addington J, Cornblatt BA, Cadenhead KS, et al. At clinical high risk for psychosis: outcome for nonconverters. *Am J Psychiatry*. 2011;168:800–805.
3. Cannon TD. Brain biomarkers of vulnerability and progression to psychosis. *Schizophr Bull*. 2016;42(suppl 1):S127–S132.
4. Sabherwal S, English JA, Föcking M, Cagney G, Cotter DR. Blood biomarker discovery in drug-free schizophrenia: the contribution of proteomics and multiplex immunoassays. *Expert Rev Proteomics*. 2016;13:1141–1155.
5. Khandaker GM, Pearson RM, Zammit S, Lewis G, Jones PB. Association of serum interleukin 6 and C-reactive protein in childhood with depression and psychosis in young adult life: a population-based longitudinal study. *JAMA Psychiatry*. 2014;71:1121–1128.
6. Chan MK, Krebs MO, Cox D, et al. Development of a blood-based molecular biomarker test for identification of schizophrenia before disease onset. *Transl Psychiatry*. 2015;5:e601.
7. van Beveren NJ, Schwarz E, Noll R, et al. Evidence for disturbed insulin and growth hormone signaling as potential risk factors in the development of schizophrenia. *Transl Psychiatry*. 2014;4:e430.
8. Schwarz E, van Beveren NJ, Ramsey J, et al. Identification of subgroups of schizophrenia patients with changes in either immune or growth factor and hormonal pathways. *Schizophr Bull*. 2014;40:787–795.
9. Perkins DO, Jeffries CD, Addington J, et al. Towards a psychosis risk blood diagnostic for persons experiencing high-risk symptoms: preliminary results from the NAPLS project. *Schizophr Bull*. 2015;41:419–428.
10. Cannon M, Clarke MC, Cotter DR. Priming the brain for psychosis: maternal inflammation during fetal development and the risk of later psychiatric disorder. *Am J Psychiatry*. 2014;171:901–905.
11. Fillman SG, Weickert TW, Lenroot RK, et al. Elevated peripheral cytokines characterize a subgroup of people with schizophrenia displaying poor verbal fluency and reduced Broca's area volume. *Mol Psychiatry*. 2016;21:1090–1098.
12. Trépanier MO, Hopperton KE, Mizrahi R, Mechawar N, Bazinet RP. Postmortem evidence of cerebral inflammation in schizophrenia: a systematic review. *Mol Psychiatry*. 2016;21:1009–1026.
13. Sekar A, Bialas AR, de Rivera H, et al. Schizophrenia risk from complex variation of complement component 4. *Nature*. 2016;530:177–183.
14. Zammit S, Kounali D, Cannon M, et al. Psychotic experiences and psychotic disorders at age 18 in relation to psychotic experiences at age 12 in a longitudinal population-based cohort study. *Am J Psychiatry*. 2013;170:742–750.
15. Drakesmith M, Dutt A, Fonville L, et al. Mediation of developmental risk factors for psychosis by white matter microstructure in young adults with psychotic experiences. *JAMA Psychiatry*. 2016;73:396–406.
16. Fonville L, Cohen Kadosh K, Drakesmith M, et al. Psychotic experiences, working memory, and the developing brain: a multimodal neuroimaging study. *Cereb Cortex*. 2015;25:4828–4838.
17. van Os J, Kenis G, Rutten BP. The environment and schizophrenia. *Nature*. 2010;468:203–212.

18. Boyd A, Golding J, Macleod J, et al. Cohort Profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2013;42:111–127.

19. Levin Y, Wang L, Schwarz E, Koethe D, Leweke FM, Bahn S. Global proteomic profiling reveals altered proteomic signature in schizophrenia serum. *Mol Psychiatry*. 2010;15:1088–1100.

20. English JA, Fan Y, Föcking M, et al. Reduced protein synthesis in schizophrenia patient-derived olfactory cells. *Transl Psychiatry*. 2015;5:e663.

21. Föcking M, Dicker P, Lopez LM, et al. Proteomic analysis of the postsynaptic density implicates synaptic function and energy pathways in bipolar disorder. *Transl Psychiatry*. 2016;6:e959.

22. Topol A, Zhu S, Hartley BJ, et al. Dysregulation of miRNA-9 in a subset of schizophrenia patient-derived neural progenitor cells. *Cell Rep*. 2016;15:1024–1036.

23. Sajic T, Liu Y, Aebersold R. Using data-independent, high-resolution mass spectrometry in protein biomarker research: perspectives and clinical applications. *Proteomics Clin Appl*. 2015;9:307–321.

24. Liu Y, Buil A, Collins BC, et al. Quantitative variability of 342 plasma proteins in a human twin population. *Mol Syst Biol*. 2015;11:786.

25. Aebersold R, Bensimon A, Collins BC, Ludwig C, Sabido E. Applications and developments in targeted proteomics: from SRM to DIA/SWATH. *Proteomics*. 2016;16:2065–2067.

26. Egertson JD, Kuehn A, Merrihew GE, et al. Multiplexed MS/MS for improved data-independent acquisition. *Nat Methods*. 2013;10:744–746.

27. Egertson JD, MacLean B, Johnson R, Xuan Y, MacCoss MJ. Multiplexed peptide analysis using data-independent acquisition and Skyline. *Nat Protoc*. 2015;10:887–903.

28. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res*. 2011;10:1794–1805.

29. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26:1367–1372.

30. Hubner NC, Bird AW, Cox J, et al. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J Cell Biol*. 2010;189:739–754.

31. Webb-Robertson BJ, Wiberg HK, Matzke MM, et al. Review, evaluation, and discussion of the challenges of missing value imputation for mass spectrometry-based label-free global proteomics. *J Proteome Res*. 2015;14:1993–2001.

32. Thorndike AN, Achtyes ED, Cather C, et al. Weight gain and 10-year cardiovascular risk with sustained tobacco abstinence in smokers with serious mental illness: a subgroup analysis of a randomized trial. *J Clin Psychiatry*. 2016;77:e320–e326.

33. Chouinard VA, Pingali SM, Chouinard G, et al. Factors associated with overweight and obesity in schizophrenia, schizoaffective and bipolar disorders. *Psychiatry Res*. 2016;237:304–310.

34. Sørensen HJ, Gamborg M, Sørensen TI, Baker JL, Mortensen EL. Childhood body mass index and risk of schizophrenia in relation to childhood age, sex and age of first contact with schizophrenia. *Eur Psychiatry*. 2016;34:64–69.

35. Strimmer K. A unified approach to false discovery rate estimation. *BMC Bioinform*. 2008;9:303.

36. Strimmer K. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics*. 2008;24:1461–1462.

37. Teo G, Kim S, Tsou CC, et al. mapDIA: preprocessing and statistical analysis of quantitative proteomics data from data independent acquisition mass spectrometry. *J Proteomics*. 2015;129:108–120.

38. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015;43:D447–D452.

39. Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke S, Neale BM, et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511:421–427.

40. Li Y, Zhou K, Zhang Z, et al. Label-free quantitative proteomic analysis reveals dysfunction of complement pathway in peripheral blood of schizophrenia patients: evidence for the immune hypothesis of schizophrenia. *Mol Biosyst*. 2012;8:2664–2671.

41. Jaros JA, Martins-de-Souza D, Rahmoune H, et al. Protein phosphorylation patterns in serum from schizophrenia patients and healthy controls. *J Proteomics*. 2012;76 Spec No.:43–55.

42. Yang Y, Wan C, Li H, et al. Altered levels of acute phase proteins in the plasma of patients with schizophrenia. *Anal Chem*. 2006;78:3571–3576.

43. Masopust J, Malý R, Andrýs C, Vališ M, Bažant J, Hosák L. Markers of thrombogenesis are activated in unmedicated patients with acute psychosis: a matched case control study. *BMC Psychiatry*. 2011;11:2.

44. Huang KC, Yang KC, Lin H, Tsao TT, Lee SA. Transcriptome alterations of mitochondrial and coagulation function in schizophrenia by cortical sequencing analysis. *BMC Genomics*. 2014;15(suppl 9):S6.

45. Hoirisch-Clapauch S, Amaral OB, Mezzasalma MA, Panizzutti R, Nardi AE. Dysfunction in the coagulation system and schizophrenia. *Transl Psychiatry*. 2016;6:e704.

46. Markiewski MM, Lambris JD. The role of complement in inflammatory diseases from behind the scenes into the spotlight. *Am J Pathol*. 2007;171:715–727.

47. Orsini F, De Blasio D, Zangari R, Zanier ER, De Simoni MG. Versatility of the complement system in neuroinflammation, neurodegeneration and brain homeostasis. *Front Cell Neurosci*. 2014;8:380.

48. Kolev M, Le Friec G, Kemper C. Complement–tapping into new sites and effector systems. *Nat Rev Immunol*. 2014;14:811–820.

49. Stephan AH, Barres BA, Stevens B. The complement system: an unexpected role in synaptic pruning during development and disease. *Annu Rev Neurosci*. 2012;35:369–389.

50. Upthegrove R, Manzanares-Teson N, Barnes NM. Cytokine function in medication-naive first episode psychosis: a systematic review and meta-analysis. *Schizophr Res*. 2014;155:101–108.

51. Curtis D. Schizophrenia genetics moves into the light. *Br J Psychiatry*. 2016;209:93–94.

52. Ruzzo EK, Geschwind DH. Schizophrenia genetics complements its mechanistic understanding. *Nat Neurosci*. 2016;19:523–525.

53. Whalley K. Psychiatric disorders: linking genetic risk to pruning. *Nat Rev Neurosci*. 2016;17:199.

54. Rutigliano G, Valmaggia L, Landi P, et al. Persistence or recurrence of non-psychotic comorbid mental disorders associated with 6-year poor functional outcomes in patients at ultra high risk for psychosis. *J Affect Disord*. 2016;203:101–110.