

ARTICLE

Open Access

Genome-wide study of immune biomarkers in cerebrospinal fluid and serum from patients with bipolar disorder and controls

Ruyue Zhang¹, Jie Song¹, Anniella Isgren², Joel Jakobsson², Kaj Blennow^{2,3}, Carl M. Sellgren^{4,5}, Henrik Zetterberg^{2,3,6,7}, Sarah E. Bergen¹ and Mikael Landén^{1,2}

Abstract

Bipolar disorder is a common, chronic psychiatric disorder. Despite high heritability, there is a paucity of identified genetic risk factors. Immune biomarkers are under more direct genetic influence than bipolar disorder. To explore the genetic associations with immune biomarker levels in cerebrospinal fluid (CSF) and blood serum which previously showed differences in bipolar disorder, we performed a study involving 291 individuals (184 bipolar disorder patients and 107 controls). The biomarkers assayed in both CSF and serum were: chitinase-3-like protein-1 (YKL-40), monocyte chemoattractant protein-1 (MCP-1), soluble cluster of differentiation (sCD14), tissue inhibitor of metalloproteinases-1 and 2 (TIMP-1 and TIMP-2). C-reactive protein (CRP) was only quantified in serum, and interleukin 8 (IL-8) measures were only available in CSF. Genome-wide association studies were conducted using PLINK for each of three genotyping waves and incorporated covariates for population substructure, age, sex, and body mass index (BMI). Results were combined by meta-analysis. Genome-wide significant associations were detected for all biomarkers except TIMP-1 and TIMP-2 in CSF. The strongest association in CSF was found for markers within the *CNTNAP5* gene with YKL-40 ($rs150248456$, $P = 2.84 \times 10^{-10}$). The strongest association in serum was also for YKL-40 but localized to the *FANCI* gene ($rs188263039$, $P = 5.80 \times 10^{-26}$). This study revealed numerous biologically plausible genetic associations with immune biomarkers in CSF and blood serum. Importantly, the genetic variants regulating immune biomarker levels in CSF and blood serum differ. These results extend our knowledge of how biomarkers showing alterations in bipolar disorder are genetically regulated.

Introduction

Bipolar disorder is a chronic psychiatric disorder characterized by recurrent episodes of mania or hypomania and depression that afflicts about 60 million people worldwide^{1–3}. The etiopathogenesis of bipolar disorder is not fully understood. Despite heritability estimates ranging between 58–93%^{4–6}, only a few dozen single

nucleotide polymorphisms (SNPs) with small effect sizes have been associated with bipolar disorder in large-scale genome-wide association studies (GWAS)^{7–11}.

One line of research has explored immunoinflammatory processes in bipolar disorder and a number of studies have investigated immune markers in serum^{12–16}. Serum immune markers are, however, not necessarily indicative of immune and inflammatory activity in the brain. This is because concentrations of cytokines and other proteins in serum or plasma come from production in peripheral tissues and thus do not reflect inflammatory processes in the brain^{17–19}. Therefore, we previously investigated a set of immune biomarkers in cerebrospinal fluid (CSF) in bipolar disorder

Correspondence: Sarah E. Bergen (sbergen@gmail.com)

¹Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

²Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden

Full list of author information is available at the end of the article
These authors contributed equally: Sarah E. Bergen, Mikael Landén

© The Author(s) 2020



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

and healthy controls, including monocyte chemoattractant protein-1 (MCP-1), chitinase-3-like protein-1 (YKL-40), soluble cluster of differentiation 14 (sCD14), tissue inhibitor of metalloproteinases-1/2 (TIMP-1 and TIMP-2), and interleukin 8 (IL-8)^{20,21}. Taken together, these results suggest that brain-specific immune mechanisms, beyond systemic inflammatory processes, are involved in the pathophysiology of bipolar disorder.

Immune biomarkers are products of gene expression and are as such measurable components in biological pathways between genotype and disease. Immune biomarkers are therefore potential endophenotypes for some psychiatric disorders by representing a measurable characteristic of the disorder more closely related to the genetic underpinnings than behavioral manifestations such as bipolar disorder^{15,22,23}. Available evidence indicates a number of loci associated with immune biomarker levels among different populations^{15,24–26}, but few genome-wide association studies of multiple immune biomarkers have been published in bipolar disorder cohorts. Revealing genetic associations for CSF and blood serum biomarkers serving as indicators of biological processes implicated in bipolar disorder might yield new insights into the genetic underpinnings of bipolar disorder.

The aim of this study was to conduct a GWAS of a set of immune biomarkers in CSF and serum that previously have been found to differ between bipolar disorder patients and controls.

Materials and methods

Subjects

The study population consisted of subjects with bipolar disorder as well as age- and sex- matched controls. Subjects were recruited from a long-term naturalistic study of bipolar disorder, the St. Göran Bipolar Project, at the bipolar outpatient unit at the Northern Stockholm psychiatric clinic in Stockholm, Sweden. The work-up and diagnostic procedures for patients and selection of controls have been described in detail previously^{20,27–29}. In brief, the key clinical assessment instrument was a Swedish version of the Affective Disorder Evaluation (ADE), which is a standardized interview protocol developed for the Systematic Treatment Enhancement Program of Bipolar Disorder (STEP-BD). The clinical diagnosis of bipolar disorder was made according to DSM-IV criteria as per the Structured Clinical Interview for DSM-IV. In addition, the Mini International Neuropsychiatric Interview (M.I.N.I.) was completed to screen for other psychiatric diagnoses. The ADE and M.I.N.I. interviews were conducted by board-certified psychiatrists, or residents in psychiatry. A best-estimate diagnostic decision was then made based on all information available by a consensus panel of experienced board-certified psychiatrists specialized in bipolar disorder. To

gather a representative cohort of bipolar disorder patients, the St. Göran study aimed to have as few exclusion criteria as possible and persons with somatic disorders including autoimmune disorder were not excluded. Patients were not remunerated for participation.

Population-based controls living in the same catchment areas were randomly selected by Statistics Sweden and contacted by mail. Details of the recruitment, and inclusion and exclusion criteria can be found elsewhere^{29,30}. Briefly, eligible persons were scheduled for a personal examination and investigated to exclude mental illness by a psychiatrist using the Mini International Neuropsychiatric Interview (M.I.N.I.) and selected parts of the ADE. Exclusion criteria were as follows: any current psychiatric disorder including personality disorder, a family history of schizophrenia or bipolar disorder in first-degree relatives, drug or alcohol abuse (based on DUDIT, AUDIT and serum levels of carbohydrate-deficient transferrin), and neurological conditions except mild migraines, as well as pregnancy, untreated endocrine disorders, dementia, and chronic systemic autoimmune disorders, except persons with controlled asthma and allergies. Control subjects were remunerated for their participation.

Only cases and controls with Scandinavian ancestry were included in this study. All participating subjects granted oral and written informed consent after complete description of the study. Ethical approvals for this study were granted by the Stockholm Regional Ethics Committee.

Sampling and biomarker analyses of CSF and blood

CSF and blood were obtained when the participants were in a stable condition. Sampling occurred between 9 and 10 am after an overnight fast. For each participant, 12 mL of CSF was collected and gently inverted to avoid gradient effects. Serum was obtained from blood samples after coagulation and centrifugation. Both serum and CSF samples were stored at -80°C at Karolinska Institutet Biobank, Sweden. The assays and kits for each biomarker have been described in detail previously^{20,21}. All biomarker concentrations were measured by experienced and board-certified laboratory technicians at the Clinical Neurochemistry Laboratory in Mölndal, Sweden who were blinded to the clinical information.

Genotyping, quality control, and imputation

Blood samples were transferred to the Karolinska Institutet Biobank for DNA extraction. An aliquot of each DNA sample was shipped to the Broad Institute (Boston, USA) for genotyping. Whole-genome genotyping was conducted at the Broad Institute using PsychChip (wave 1), Affymetrix 6.0 (wave 2), and Illumina OmniExpress (wave 3) chips. All controls were in wave 1 and wave 3,

while the bipolar cases were exclusively in wave 2. Following quality control steps, datasets were then imputed with the full 1000 Genomes Project integrated variant set as reference³¹. Details of genotyping, quality control, and imputation procedures have been described previously³². The three waves shared more than 10 million SNPs after imputation.

A total of 183 bipolar cases and 107 controls from the St. Görán project had both biomarker information and genotype data available.

Statistical analyses

We used PLINK version 1.9, SPSS Statistics 23 and R version 3.3.3 with the qqman package for all statistical analyses. Group differences between bipolar cases and controls were tested using the Mann-Whitney U test for age, BMI, and all biomarker concentrations, while Fisher's exact test was used for sex. Interquartile range was calculated for continuous variables.

The CRP distributions deviated from normality, prohibiting use of linear regressions in GWAS analysis. We therefore conducted a log transformation to adapt it to a distribution compatible with linear regression.

The GWAS analysis was conducted using imputed SNP dosages and linear regression models in PLINK for each wave and incorporated the first four multi-dimensional scaling (MDS) components, age, sex, and BMI as covariates. Results were then combined by meta-analysis in PLINK using a random-effects model. Although imputed data were used, there were still SNPs that were not present in one or more waves. Waves included in the meta-analysis are noted in results. SNPs with minor allele frequency (<1%) and poor imputation (INFO < 0.6) were removed from meta-analysis results. A standard genome-wide significance threshold of $p < 5 \times 10^{-8}$ was used³³. Linkage disequilibrium based clumping was used for grouping correlated SNPs to define regions of association. Genes were also identified based on UCSC hg19 coordinates. GWAS summary statistics were also uploaded to FUMA to obtain results from MAGMA gene-set analysis and MAGMA tissue expression analysis (GTEx v6, 30 general tissue types)^{34,35}.

SNPs with p -values $< 5 \times 10^{-5}$ from either CSF or serum results were used for the difference test to compare the effect sizes between CSF and serum results. Effect sizes of these SNPs were standardized by dividing by the standard deviations of the biomarker levels. Pearson correlation coefficients were used to test the correlation of standardized effect sizes from CSF and serum results.

Code availability

The scripts used to generate these results can be provided upon request.

Results

Demographics and clinical characteristics

We included 184 bipolar disorder patients (67 men and 117 women) and 107 controls (46 men and 61 women) for the serum measurements. A subset of this sample population—114 bipolar disorder patients (44 men and 70 women) and 83 controls (36 men and 47 women)—who consented to lumbar puncture comprised the study population for CSF measurements. A summary of sample sizes included in this study for each genotyping wave is described in Supplementary Table S1.

A description of demographic and clinical characteristics of the study population as well as comparison of differences between bipolar patients and controls can be found in Supplementary Table S2. Note that the comparisons of biomarker concentrations between cases and controls shown in Table S2 have been published previously^{20,21}, but may differ somewhat as not all individuals had both biomarker information and genotype data available.

Genetic variants associated with immune biomarkers in CSF

A total of >5.6 million SNPs from 114 bipolar patients and 83 controls were included in our final meta-analysis for biomarkers in CSF. The quantile-quantile plots for each of the six biomarkers showed moderate deviations from the null distribution at low p -values indicating the presence of association signals, but no deviation at the higher p -values, which denotes well matched cases and controls, i.e., no inflation (Fig. S1). The genomic inflation factor, λ , ranged from 1.05–1.11.

The genome-wide association analysis results for the six biomarkers in CSF are shown in Fig. 1. The number of genome-wide significant regions (Index SNPs' P -values $< 5 \times 10^{-8}$) for each GWAS meta-analysis on YKL-40, MCP-1, sCD14, and IL-8 in CSF was 4, 3, 2, and 6, respectively. No SNPs of genome-wide significance were associated with TIMP-1 and TIMP-2.

Table 1 illustrates the genome-wide significant (GWS) regions for each biomarker in CSF. The top GWS SNPs associated with YKL-40 were located in the genes *CNTNAP5*, *EYS*, and *FER1L3*. For MCP-1, the SNP with the strongest association was located within *ACAA2* gene (rs10438979, $P_{A/G} = 1.64 \times 10^{-9}$). The other two top GWS SNPs were located near or within *LINCO1288*, *FAM129A*, and *EDEM3*. Two GWS SNPs were related to sCD14 levels in CSF, which are located in the gene area of *C8orf37-AS1* and *TUT1*.

The GWS SNPs associated with IL-8 were located in the genes *ZNF876P*, *LINC01192*, *SEPT2*, *FARP2*, and *SATB2*.

The results from MAGMA tissue expression analysis for 30 general tissue types for biomarkers from CSF are shown in Fig. S2. The genetic variants associated with

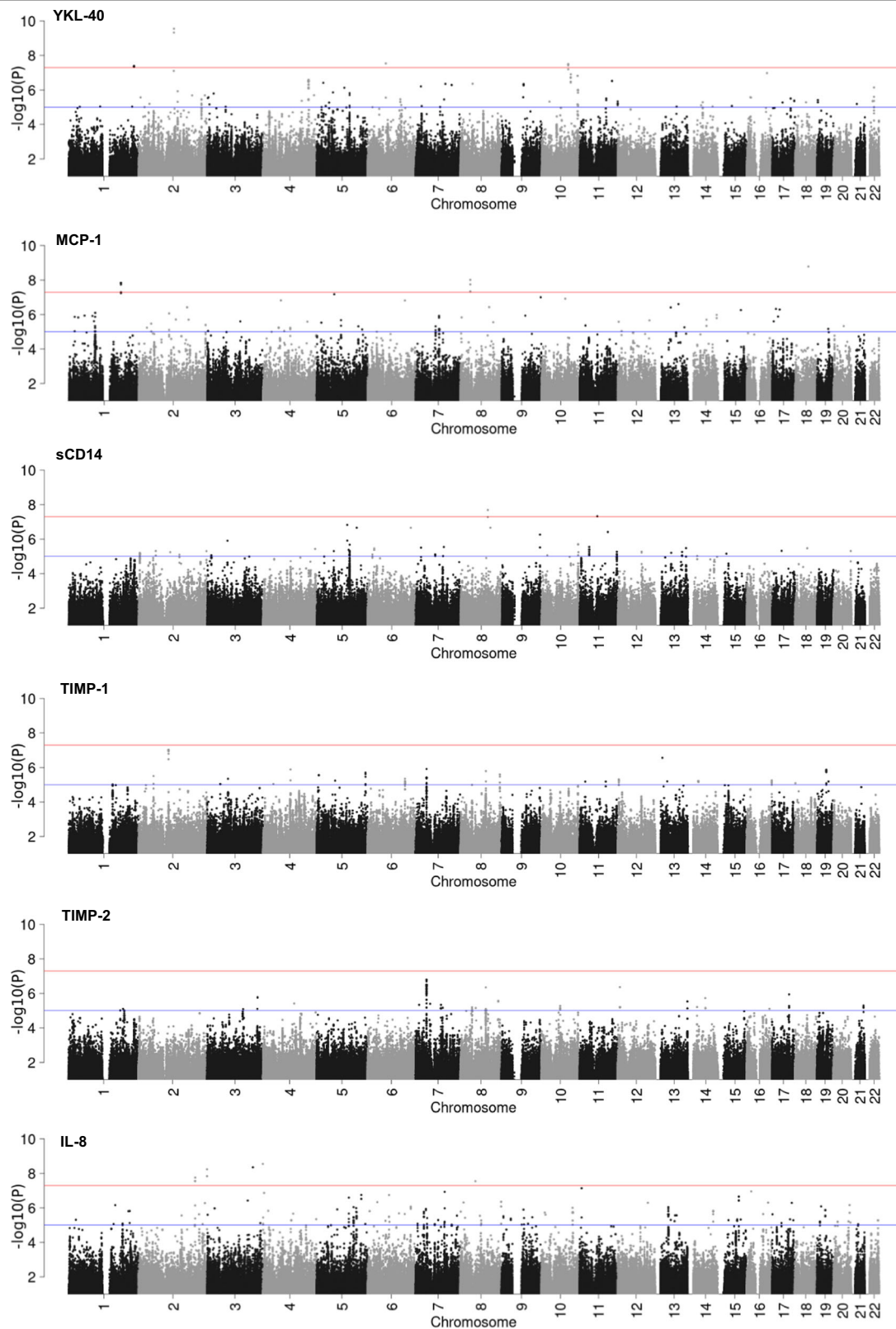


Fig. 1 Manhattan plots for six CSF immune biomarkers. For each of the six biomarkers, a Manhattan plot shows the $-\log_{10} P$ -value (vertical axis) for all SNPs that pass the standard quality control filters, have a minor allele frequency $> 1\%$ and imputation info score > 0.6 along with the chromosome position (horizontal axis). Chromosomes are shown in alternating colors for clarity. The red lines denote genome-wide significant P -values $< 5 \times 10^{-8}$ while the blue lines show suggestive associations with P -values $< 1 \times 10^{-5}$.

Table 1 Top GWS regions for each biomarker in CSF.

Chr	Index SNP	A1/A2	Freq	β	p -value	N	Position	KB	Genes	Waves
<i>YKL-40</i>										
2	rs150248456	A/G	0.989	-88.276	2.84e-10	2	chr2:125287495..125383005	95.51	CNTNAP5	1,3
6	rs11753319 ^a	C/G	0.951	-39.387	2.89e-08	2	chr6:64902720..65013122	110.40	EYS	1,2,3
10	rs182911102 ^a	T/C	0.985	-64.832	3.15e-08	4	chr10:94991992..95075604	83.61	FER1L3	1,2,3
1	rs16856846	A/T	0.020	62.660	4.11e-08	3	chr1:232470911..232485096	14.19		1,2
<i>MCP-1</i>										
18	rs10438979 ^a	A/G	0.960	-230.358	1.64e-09	1	47267946	0	ACAA2	1,2,3
8	rs193114076 ^a	A/G	0.011	370.998	9.68e-09	3	chr8:34726827..34789467	62.64	LINC01288	1,2,3
1	rs114271760	A/G	0.014	362.410	1.45e-08	5	chr1:184697347..184966318	268.97	FAM129A,EDEM3	1,3
<i>sCD14</i>										
8	rs182557857	T/C	0.012	94.392	2.05e-08	4	chr8:96576432..96626315	49.88	C8orf37-AS1	1,3
11	rs117456286	A/G	0.013	89.513	4.76e-08	1	62345761	0	TUT1	1,3
<i>IL-8</i>										
4	rs142700748 ^a	T/C	0.975	-20.461	2.82e-09	1	234362	0	ZNF876P	2,3
3	rs184700722	T/C	0.011	-16.515	4.41e-09	1	162966687	0	LINC01192	1,3
2	rs182193570 ^a	A/G	0.017	26.143	5.79e-09	1	242292698	0	SEPT2	2,3
2	rs187443429 ^a	T/C	0.024	23.725	1.45e-08	1	242392862	0	FARP2	2,3
2	rs9646754 ^a	T/C	0.034	20.001	1.74e-08	1	200300101	0	SATB2	1,2,3
2	rs115307848 ^a	A/G	0.021	22.536	2.79e-08	3	chr2:199894134..200004958	110.82		1,2,3

Genes in bold include the index SNP. Genes not in bold are the gene area that SNPs are closest to. GWS is defined as $P < 5e-08$. A1/A2, beta (β), p -value, waves are based on the meta-analysis of GWAS data from each wave. Position is the basepair position or given in UCSC hg19 coordinates.

Chr chromosome, (Index)SNP single nucleotide polymorphism (with the strongest association in the genomic region), A1/A2 reference and alternate allele, Freq weighted average frequency of reference allele, β random-effects meta-analysis β estimate, p -value random-effects meta-analysis p -value, N number of SNPs in the reported region, Waves valid waves included for the SNP.

^aResults from meta-analysis including wave 2.

YKL-40 from CSF were significantly expressed in bladder. No other significant findings from GTEx were identified.

Genetic variants associated with serum immune biomarkers

Over five million SNPs from 184 bipolar patients and 107 controls were included in our final meta-analysis for biomarkers in serum. The quantile-quantile plots for each of the six biomarkers were consistent with no inflation and strong association signals (Fig. S3). The genomic inflation factor, λ , ranged from 1.01 to 1.05.

The genome-wide association analysis results for the six biomarkers in serum are illustrated in Fig. 2. The serum biomarkers (and number of GWS regions) were: YKL-40 ($n = 37$), MCP-1 ($n = 1$), sCD14 ($n = 2$), TIMP-1 ($n = 1$), TIMP-2 ($n = 4$), and CRP ($n = 1$).

Table 2 shows the GWS regions for each biomarker in blood serum. Three top GWS SNPs associated with YKL-40 levels map to the gene *FANCI*. One SNP reached GWS for MCP-1 (rs34625674, $P_{D/18} = 1.70 \times 10^{-8}$) and one for TIMP-1 (rs118161330, $P_{A/G} = 2.50 \times 10^{-8}$). The top GWS

SNPs associated with sCD14 were located in the gene *SLC8A1*. Notably, after meta-analysis with all three waves, the largest number of SNPs ($n = 110$) associated with sCD14 were identified on chromosome 5q31.3 (top: rs2569191, $P_{T/C} = 7.09 \times 10^{-9}$, $\beta = -125.703$, $MAF = 0.421$). More than 15 genes are located in this area. One gene-set remained significantly associated with sCD14 after Bonferroni correction (Table S3). As for TIMP-2, the largest number of SNPs ($n = 36$) were identified near the gene area of *LOC101927410* (top: rs150356358, $P_{A/G} = 2.16 \times 10^{-8}$, $\beta = -13.907$, $MAF = 0.211$). The lone GWS SNP associated with log (CRP) was located within the *LIBASH3B* gene.

The results from MAGMA tissue expression analysis for 30 general tissue types for biomarkers from serum are shown in Fig. S4. None demonstrated significance.

Difference tests for top SNPs in CSF and serum

The genome-wide significant SNPs for serum and CSF for each biomarker with both measures did not overlap. To formally test for similarity in sub-significance

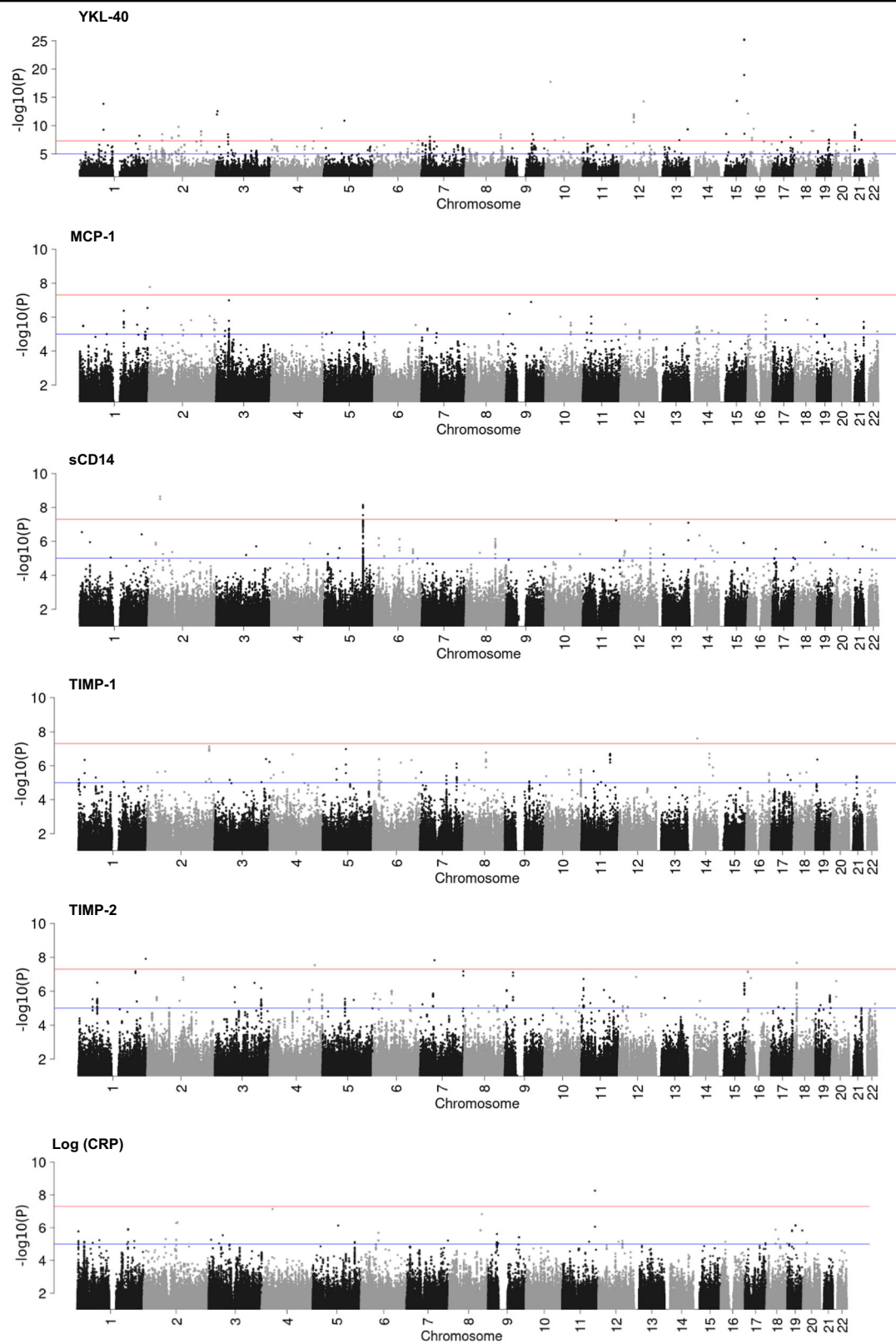


Fig. 2 Manhattan plots for six serum immune biomarkers. For each of the six biomarkers, a Manhattan plot shows the $-\log_{10} P$ -value (vertical axis) for SNPs that pass the standard quality control filters, have a minor allele frequency $> 1\%$ and imputation info score > 0.6 along with the chromosome position (horizontal axis). Chromosomes are shown in alternating colors for clarity. The red lines denote genome-wide significant P -values $< 5 \times 10^{-8}$ while the blue lines show suggestive associations with P -values $< 1 \times 10^{-5}$.

Table 2 Top GWS SNPs for each immune biomarker in blood serum.

Chr	Index SNP	A1/A2	Freq	β	p-value	N	Position	KB	Genes	Waves
YKL-40										
15	rs188263039	T/G	0.988	-76.738	5.80e-26	1	89809787	0	FANCI	1,3
15	rs117809422	A/G	0.015	76.377	7.24e-26	1	89815928	0	FANCI	1,3
15	rs188631178	T/C	0.986	-67.032	1.15e-19	1	89845859	0	FANCI	1,3
10	rs17602253 ^a	T/C	0.988	-65.955	1.79e-18	1	18334915	0	SLC39A12	1,2
15	rs62012778 ^a	A/C	0.966	-37.325	4.41e-15	4	chr15:63818629..63895362	76.73	USP3,FBXL22	1,2,3
1	rs116057745 ^a	A/G	0.982	-31.092	1.47e-14	1	85746632	0		1,2,3
16	rs117868506 ^a	T/C	0.013	50.578	7.44e-13	2	chr16:1420638..1428236	7.60	UNKL	2,3
12	rs117843845 ^a	T/C	0.020	47.572	9.12e-13	8	chr12:48287373..48314900	27.527	VDR	1,2
3	rs182265848 ^a	T/G	0.014	58.516	1.05e-12	1	chr3:2579507..2579507	0	CNTN4	2,3
5	rs35017964 ^a	T/C	0.989	-48.149	1.33e-11	1	chr5:73222206..73222206	0		2,3
2	rs190160997 ^a	A/C	0.011	45.776	1.58e-10	4	chr2:107295383..107436962	141.579	ST6GAL2	2,3
4	rs116622969 ^a	T/C	0.983	-36.776	2.80e-10	1	chr4:182066384..182066384	0		1,2,3
16	rs112260868	A/G	0.012	53.408	3.50e-10	1	chr16:21613268..21613268	0	METTL9	1,3
13	rs150949191 ^a	A/C	0.015	37.114	4.56e-10	2	chr13:109235412..109238352	2.94		2,3
1	rs114751860	T/C	0.013	40.152	5.25e-10	1	chr1:85883897..85883897	0	DDAH1	1,3
18	rs172586 ^a	A/G	0.020	46.598	8.46e-10	1	chr18:60321603..60321603	0		2,3
18	rs512482 ^a	T/G	0.978	-24.412	8.81e-10	1	chr18:66119026..66119026	0		2,3
2	rs186493621 ^a	A/C	0.985	-36.147	9.29e-10	16	chr2:187931545..188399880	468.34	TFPI,CALCRL	2,3
21	rs28707003 ^a	T/C	0.958	-28.903	1.33e-09	12	chr21:15202015..15222754	20.74		1,2,3
15	rs67659323 ^a	C/G	0.960	-29.117	2.98e-09	1	chr15:25490757..25490757	0		1,2,3
9	rs117574354	C/G	0.015	25.396	3.08e-09	5	chr9:94685768..94923946	238.178	SPTLC1,ROR2	1,3
2	rs150886752 ^a	T/C	0.017	40.299	3.36e-09	2	chr2:47882212..47901294	19.08		1,2,3
3	rs34307719 ^a	T/C	0.987	-45.440	3.38e-09	1	chr3:42943216..42943216	0		1,2
8	rs960667 ^a	T/G	0.976	-24.890	3.69e-09	5	chr8:126632214..126636054	3.84		1,2
2	rs116200463 ^a	T/C	0.013	40.224	6.12e-09	3	chr2:107460348..107488170	27.82	ST6GAL2	2,3
1	rs2290199 ^a	A/G	0.035	20.132	6.21e-09	3	chr1:214825751..214853690	27.94	CENPF	1,2,3
7	rs6969115 ^a	A/G	0.018	35.247	9.34e-09	18	chr7:30250241..30281728	31.49		1,2
17	rs4791291 ^a	T/C	0.924	-16.959	1.10e-08	4	chr17:65369397..65396158	26.76	PITPNC1	1,2,3
3	rs192944990 ^a	T/C	0.012	40.733	1.14e-08	2	chr3:43867670..43888114	20.44		1,2
16	rs138053699 ^a	A/G	0.019	46.309	1.40e-08	2	chr16:15812581..15821114	8.53	NDE1,MYH11	1,2
2	rs58011745 ^a	A/G	0.035	25.632	1.68e-08	1	chr2:82922195..82922195	0		1,2,3
4	rs116608013 ^a	A/C	0.987	-49.182	2.76e-08	1	chr4:2397776..2397776	0	ZFYVE28	1,2
9	rs112153762 ^a	C/G	0.977	-39.896	3.10e-08	4	chr9:98468727..98480513	11.79		2,3
19	rs118145918 ^a	T/G	0.024	32.162	3.13e-08	17	chr19:44419463..44487889	68.43	ZNF45,ZNF221	1,2
10	rs113755171 ^a	T/C	0.011	45.313	3.47e-08	1	chr10:34100343..34100343	0		2,3
6	rs115196122 ^a	A/G	0.983	-41.511	4.36e-08	3	chr6:159346245..159354145	7.9		1,2
7	rs11979716 ^a	T/G	0.022	32.630	4.57e-08	8	chr7:30292334..30320400	28.07		1,2

Table 2 continued

Chr	Index SNP	A1/A2	Freq	β	p -value	N	Position	KB	Genes	Waves
<i>MCP-1</i>										
2	chr2_3403609_D ^a	D/18	0.033	254.836	1.70e-08	1	3403609	0	TRAPPC12	1,2,3
<i>sCD14</i>										
2	rs190197089 ^a	T/C	0.015	700.006	2.27e-09	2	chr2:40635776..40659665	0	SLC8A1	1,2
5	rs2569191 ^a	T/C	0.579	-125.703	7.09e-09	110	139779258..140186438	407.18	ZMAT2,WDR55,TMCO6,SRA1,SLC35A4,PCDHA2,PCDHA3,PCDHA1,NDUFA2,IK,HARS,HARS2,EIF4EBP3,DND1,CD14,APBB3,ANKHD1,ANKHD1-EIF4EBP3	1,2,3
<i>TIMP-1</i>										
14	rs118161330	A/G	0.017	-140.235	2.50e-08	1	33801149	0	NPAS3	1,3
<i>TIMP-2</i>										
1	chr1_243283251_D ^a	D/16	0.018	54.173	1.24e-08	1	243283251	0		1,2,3
7	rs117145170 ^a	A/T	0.019	60.350	1.51e-08	1	52273737	0		1,2
18	rs150356358 ^a	A/G	0.789	-13.907	2.16e-08	36	10640992..10656164	15	<i>LOC101927410</i>	1,2,3
4	rs114147769 ^a	A/T	0.968	-42.516	2.94e-08	1	162690147	0	FSTL5	2,3
<i>log(CRP)</i>										
11	rs57213254	T/C	0.011	3.16	5.51e-09	2	chr11:122678513..122679916	1.40	UBASH3B	1,3

Bold gene names include the index SNP within the gene. Genes not in bold are those that SNPs are closest to. GWS is defined as $P < 5e-08$. A1/A2, beta(β), p -value, waves are based on the meta-analysis of GWAS data from each wave. Position is the basepair position or given in UCSC hg19 coordinates.

Chr chromosome, (Index)SNP single nucleotide polymorphism (with the strongest association in the genomic region), A1/A2 reference and alternate allele, Freq weighted average frequency of reference allele, β random-effects meta-analysis β estimate, p -value random-effects meta-analysis p -value, N number of SNPs in the reported region, Waves valid waves included for the SNP.

^aResults from meta-analysis including wave 2.

threshold results, we compared the difference in effect sizes for top SNPs in both measures using a correlation test. The correlation coefficients of each biomarker pair and corresponding p -values are presented in Table 3.

From this test, there were significant positive correlations between standardized effect sizes of top SNPs in different tissues in YKL-40 ($r = 0.533$, $P < 0.001$), MCP-1 ($r = 0.449$, $P < 0.001$), and sCD14 ($r = 0.092$, $P = 0.020$). However, the correlation in sCD14 was rather modest while the standardized effect sizes of top SNPs from different tissues in YKL-40 and MCP-1 showed moderate correlation. The correlation coefficients of TIMP-1 ($r = -0.036$, $P = 0.360$) and TIMP-2 ($r = -0.008$, $P = 0.796$) were close to zero, which indicates no relationship between effect sizes of SNPs associated with CSF or serum.

Discussion

This is the first genome-wide study of multiple immune biomarkers in serum and CSF in bipolar disorder. Despite the modest sample size, a number of SNPs reached genome-wide statistical significance.

Table 3 Pearson correlation coefficients (r) between standardized effect sizes of top SNPs from CSF and serum results.

Biomarkers	YKL-40	MCP-1	sCD14	TIMP-1	TIMP-2
r	0.533	0.449	0.092	-0.036	-0.008
P -value	<0.001*	<0.001*	0.020*	0.360	0.796
N	937	444	638	658	963

N number of SNPs included in the test.

* P -value < 0.05.

CSF markers

Among immune biomarkers in CSF, the strongest association was found between rs150248456 and YKL-40. Rs150248456 is located within an intron of *CNTNAP5* (contactin associated protein like 5) on chromosome 2. The product of *CNTNAP5* belongs to the neurexin family³⁶. Interestingly, SNPs located in *CNTNAP5* have been found to be significantly associated with mathematical ability, self-reported educational attainment, cognitive performance and response to antipsychotic treatment in

schizophrenia^{37,38}. Two GWS SNPs—only represented in genotyping waves containing control subjects—were observed in this area. Although CSF concentration of YKL-40 differs between bipolar cases and controls²⁰, genes are likely to regulate the expression of biomarkers in people with and without bipolar disorder in a similar way. Interestingly, a previous GWAS of bipolar disorder in Norwegian individuals followed by replication in Icelandic samples also found nominally significant ($P < 0.05$) markers located in this gene area³⁹. The top significant region associated with CSF YKL-40 in bipolar disorder cases was in the gene *EYS* (rs11753319), which is a novel locus. SNPs located in *EYS* were also reported to be significantly associated with mathematical ability, self-reported educational attainment, BMI, alcohol drinking, systolic blood pressure, mood disorder, unipolar depression, and schizophrenia^{37,40–43}. Several SNPs located near *LINC01288* and *EDEM3* were found to be genome-wide significantly associated with MCP-1. *LINC01288* was reported to be significantly associated with BMI, while *EDEM3* was found to be associated with educational attainment, mathematical ability, systemic lupus erythematosus, and cognitive performance^{37,44,45}. Four GWS SNPs located in *C8orf37-AS1* were found to be associated with CSF concentration of sCD14. Heel bone mineral density and smoking behaviors have previously been associated with *C8orf37-AS1*.

Although no SNPs reached genome-wide significance for TIMP-1 or TIMP-2 concentration in CSF, several association peaks for the two biomarkers nearly reached GWS (Fig. 2). This suggests that larger sample sizes may yield significant associations.

Among genes associated with IL-8 in CSF, *SATB2* is highly expressed in brain⁴⁶, indicating a potential molecular mechanism from associated SNPs to IL-8 CSF levels.

Serum markers

As for immune biomarkers in serum, the top three GWS SNPs associated with YKL-40 are located in the intron area of gene *FANCI* on chromosome 15. The product of the *FANCI* gene is a member of the Fanconi anaemia complementation (FANC) group⁴⁶. Genetic variation in *FANC* group has previously been found to be associated with psychiatric illness⁴⁷. When considering both cases and controls, the top SNP is located near the gene *SLC39A12*. This gene is highly expressed in brain and was identified by bioinformatic analyses as a significant molecular biomarker in the progression of psychiatric disorders, including bipolar disorder⁴⁸. Moreover, the genes *SPTLC1* and *ROR2* that were associated with YKL-40 level in serum are moderately associated with schizophrenia and bipolar disorder ($P = 5.79 \times 10^{-7}$) according to a previous GWAS⁴⁹.

The strongest association with sCD14 in serum was with rs190197089. This SNP is located within an intron of

SLC8A1 (solute carrier family 8 member A1) on chromosome 2. There is another GWS SNP located in the same area. Both of them are from the meta-analysis of two waves (wave 1 and wave 2), but were unavailable in wave 3. Interestingly, the *SLC8A1* gene has previously been associated with bipolar disorder^{50,51}. The largest number of SNPs ($n = 110$) associated with sCD14 after meta-analysis with all 3 waves were identified on chromosome 5q31.3 (top: rs2569191). More than 15 genes are located in this area. The 5q31.3 region has previously been linked to bipolar disorder⁵⁰. The other top SNP associated with sCD14 in serum, rs2569191, is located near the *CD14* gene. *CD14* encodes a surface antigen that is expressed by monocytes⁴⁶.

There was one GWS locus associated with TIMP-1 level in serum. It is located near *NPAS3*, which is a transcription factor involved in neurogenesis. *NPAS3* has been implicated in a pathway associated with bipolar disorder⁵². However, this was the result only from controls. Among genes where SNPs associated with TIMP-2 were located, *FSTL5* is the gene with the highest expression in brain among all tissues⁵³.

After log transformation of CRP levels in serum, one SNP (rs57213254) was GWS. Altered CRP levels in serum have been found to be associated with not only bipolar disorder, but also hypertension, stroke, coronary heart disease, type 2 diabetes mellitus, and cancer^{54–58}. A number of loci were associated with serum CRP levels in a previous GWAS in the general population²⁵. However, the GWS SNP associated with CRP in this study is novel. Replication is needed to confirm the association between this GWS SNP and altered CRP levels in bipolar patients and controls.

Compared with previous studies in larger samples, we still found a number of novel GWS SNPs associated with disease-associated immune biomarker level differences. This might imply possible biological pathways from genetic markers to inflammatory mediators in bipolar disorder. However, additional work is needed to confirm this conjecture. There are several potential explanations for the new associations despite the small sample size: First, previous GWAS of biomarkers have studied the general population or people with Alzheimer's disease^{24,25}. The novel GWS SNPs found in the meta-analysis including wave 2 (which was case-only) might be associated with biomarker differences in bipolar disorder specifically. Second, although imputation was conducted on each wave, some SNPs did not exist across all three waves due to array differences or fluctuations in allele frequencies across waves. The GWS results from only the control waves which were not significant in prior studies^{25,59} have very low minor allele frequencies (MAF < 0.015). They might therefore have been removed from wave 2 during quality control steps and may not have been analyzed in other

studies. However, the biomarker levels differ between bipolar cases and controls, and genes regulate the expression of biomarkers in people with and without bipolar disorder in a similar way²⁰. We could also posit that these findings might be specific to the Swedish population. Further studies are needed to affirm these assertions. Finally, it is possible that the slight deviations from normal distributions for biomarkers other than CRP could have resulted in false positive associations. However, the statistical tests used are robust to minor deviations in normality. False positives are a concern for all studies and may lead to variation in significant SNPs across studies, underscoring the importance of replication efforts.

The number of significant SNPs associated with disease-associated immune biomarkers differed depending on whether the sampling substrate was CSF or serum. GWS SNPs associated with blood serum biomarkers are more numerous than GWS SNPs in relation to CSF biomarkers. The reason might be the smaller sample size for CSF than blood serum. The associated SNPs and the genes where the SNPs localized are different between the same biomarker in CSF and serum. However, when comparing nominally associated SNPs from CSF and serum, effect sizes for YKL-40 and MCP-1 from CSF and serum are moderately correlated. Little correlation was observed in sCD14 while no correlation was found for TIMP-1 and TIMP-2. This mirrors findings from a previous study in this sample which observed correlations between CSF and serum levels for YKL-40 and MCP-1²⁰ but not for TIMP-1 and TIMP-2. These convergent lines of evidence indicate that some biomarkers share genetic regulation across tissues while others do not.

There are limitations of this study to consider. First, given the hurdle of lumbar puncture to collect CSF, the sample size is limited and we were unable to identify a replication cohort. Despite this, we found numerous loci associated with immune biomarkers, but our sample size is almost certainly not sufficient to identify all relevant loci. Second, genotyping was conducted on different chips (PsychChip, Affymetrix 6.0, and Illumina OmniExpress) that provide incomplete overlap of directly genotyped markers. Although we use the imputed SNP dosages in order to increase the number of same SNPs shared by all three waves, the overlap is incomplete. Finally, there might be confounders associated with bipolar disorder such as smoking and alcohol use. Smoking and alcohol abuse might not only affect immune-related biomarker levels, but are also partly genetically mediated^{60,61}. It is therefore possible that some significant genetic associations are due to the intermediary effect of smoking or alcohol consumption rather directly linked to biomarkers levels.

Despite the limitations, there are several strengths of this study. First, this is the first combined GWAS of peripheral and central immune biomarkers focusing on

bipolar disorder to date, which shed light on the genetic regulation of the immune system in bipolar disorder. Second, we studied biomarkers measured in CSF that closely reflect the chemistry of the brain. Third, there is no evidence of heterogeneity for markers with the strongest association in our study, using Q-tests ($P > 0.50$) and I^2 index ($I^2 = 0.00$), while other SNPs showed high heterogeneity with Q-tests ($P < 0.10$) and I^2 index ($I^2 > 10$). We used the random-effects model for meta-analysis, which conservatively accounts for heterogeneity. We detected a number of SNPs associated with immune biomarker levels with large effect sizes. Compared to results from previous GWAS on bipolar disorder, the strength of the associations (i.e., effect sizes) between genetic variations and immune mediators is stronger than that with bipolar disorder itself^{8,11,39}.

This study raises the possibility of using specific genetic markers as a proxy for immune biomarkers measured in CSF, which would be less cumbersome than collecting CSF through lumbar puncture. As sample size increases, the predictive ability will continue to improve. If replicated, the genetic markers may thus serve as indicators of biological processes implicated in psychiatric disorders. Indeed, specific genetic markers related to immune biomarkers profiles might be used to differentiate between different psychiatric diseases.

In summary, a number of biologically plausible SNPs significantly influencing immune biomarker levels in CSF and serum which demonstrated prior alterations in bipolar disorder have been identified in this study. Several of these SNPs are located in genes reported to be associated with bipolar disorder. The genetic variants associated with immune biomarker levels in CSF differ compared with those in serum. However, nominally significant variants showed correlated effect sizes for some biomarkers and this generally corresponded to whether the biomarker levels directly exhibited correlations between CSF and serum. The results of these GWAS can provide a route for the future investigations of genetic factors and immune biomarkers to aid in accurate diagnosis and development of treatments for bipolar disorder.

Acknowledgements

We would like to thank all the individuals who participated in this study and the staff at the St. Göran bipolar affective disorder unit, including coordinator Martina Wennberg, study nurses Agneta Carlswård-Kjellin, Lena Lundberg, and Benita Gezelius, and data managers Haydeh Olofsson and Mathias Kardell. Yngve Hallström is acknowledged for performing lumbar punctures. We finally wish to thank the BBMRI.se and KI Biobank at Karolinska Institutet for professional biobank service. This research was supported by grants from the Swedish Research Council (2018-02653, 2018-02532, and 2017-00915), the Swedish foundation for Strategic Research (KF10-0039), the Swedish Brain foundation (#FO2017-0243), and the Swedish Federal Government under the LUA/ALF agreement (ALF 20170019, ALFGBG-716801, ALFGBG-715986, and ALFGBG-720931). H.Z. is a Wallenberg Academy Fellow. C.S.M. is a Wallenberg Clinical Fellow. K.B. holds the Torsten Söderberg Professorship in Medicine at the Royal Swedish Academy of Sciences.

Author details

¹Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. ²Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, the Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden. ³Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden. ⁴Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden. ⁵Centre for Psychiatry Research, Karolinska Institutet, & Stockholm Health Care Services, Stockholm County Council, Karolinska University Hospital, Stockholm, Sweden. ⁶Department of Neurodegenerative Disease, UCL Queen Square Institute of Neurology, Queen Square, London, UK. ⁷UK Dementia Research Institute at UCL, London, UK

Conflict of interest

H.Z. has served at scientific advisory boards for Roche Diagnostics, Wave, Samumed and CogRx, has given lectures in symposia sponsored by Biogen and Alzecure, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB, a GU Ventures-based platform company at the University of Gothenburg (all unrelated to the submitted work). J.J. has been an employee of AstraZeneca, Apr 2017-Jan 2019 (not related to the submitted work). K.B. has served as a consultant or at advisory boards for Alector, Alzheon, CogRx, Biogen, Lilly, Novartis and Roche Diagnostics, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB, a GU Venture-based platform company at the University of Gothenburg, all unrelated to the work presented in this paper. C.S.M. is a co-founder of Outermost Therapeutics (unrelated to the submitted work). The other authors declare no conflict of interest.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41398-020-0737-6>).

Received: 11 August 2019 Revised: 13 January 2020 Accepted: 20 January 2020

Published online: 05 February 2020

References

- Judd, L. L. et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Arch. Gen. Psychiatry* **59**, 530–537 (2002).
- Judd, L. L. et al. A prospective investigation of the natural history of the long-term weekly symptomatic status of bipolar II disorder. *Arch. Gen. Psychiatry* **60**, 261–269 (2003).
- World Health Organization. Mental disorders [Internet]. <http://www.who.int/mediacentre/factsheets/fs396/en/> (2017).
- McGuffin, P. et al. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch. Gen. Psychiatry* **60**, 497–502 (2003).
- Song, J. et al. Bipolar disorder and its relation to major psychiatric disorders: a family-based study in the Swedish population. *Bipolar Disord.* **17**, 184–193 (2015).
- Kieseppa, T., Partonen, T., Haukka, J., Kaprio, J. & Lonnqvist, J. High concordance of bipolar I disorder in a nationwide sample of twins. *Am. J. Psychiatry* **161**, 1814–1821 (2004).
- Muhleisen, T. W. et al. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat. Commun.* **5**, 3339 (2014).
- Baum, A. E. et al. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol. Psychiatry* **13**, 197–207 (2008).
- Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
- Soronen, P. et al. Replication of GWAS of bipolar disorder: association of SNPs near CDH7 with bipolar disorder and visual processing. *Mol. Psychiatry* **15**, 4–6 (2010).
- Klar, P. et al. Whole-genome association study of bipolar disorder. *Mol. Psychiatry* **13**, 558–569 (2008).
- Rosenblat, J. D. & McIntyre, R. S. Bipolar disorder and immune dysfunction: epidemiological findings, proposed pathophysiology and clinical implications. *Brain Sci.* **7**, 144 (2017).
- Muneer, A. Bipolar disorder: role of inflammation and the development of disease biomarkers. *Psychiatry Investig.* **13**, 18–33 (2016).
- Cremschi, L. et al. Prevalences of autoimmune diseases in schizophrenia, bipolar I and II disorder, and controls. *Psychiatry Res.* **258**, 9–14 (2017).
- Sellgren, C. M. et al. A genome-wide association study of kynurenic acid in cerebrospinal fluid: implications for psychosis and cognitive impairment in bipolar disorder. *Mol. Psychiatry* **21**, 1342–1350 (2016).
- Olsson, S. K., Sellgren, C., Engberg, G., Landén, M. & Erhardt, S. Cerebrospinal fluid kynurenic acid is associated with manic and psychotic features in patients with bipolar I disorder. *Bipolar Disord.* **14**, 719–726 (2012).
- Bromander, S. et al. Changes in serum and cerebrospinal fluid cytokines in response to non-neurological surgery: an observational study. *J. Neuroinflammation* **9**, 242 (2012).
- Maier, B., Laurer, H. L., Rose, S., Buuman, W. A. & Marzi, I. Physiological levels of pro- and anti-inflammatory mediators in cerebrospinal fluid and plasma: a normative study. *J. Neurotrauma* **22**, 822–835 (2005).
- Sellgren, C. M. et al. Peripheral and central levels of kynurenic acid in bipolar disorder subjects and healthy controls. *Transl. Psychiatry* **9**, 37 (2019).
- Jakobsson, J. et al. Monocyte and microglial activation in patients with mood-stabilized bipolar disorder. *J. Psychiatry Neurosci.* **40**, 250–258 (2015).
- Isgren, A. et al. Increased cerebrospinal fluid interleukin-8 in bipolar disorder patients associated with lithium and antipsychotic treatment. *Brain Behav. Immun.* **43**, 198–204 (2015).
- Gottesman, I. & Gould, T. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am. J. Psychiatry* **160**, 636–645 (2003).
- Sigitova, E., Fisar, Z., Hroudova, J., Cikanova, T. & Raboch, J. Biological hypotheses and biomarkers of bipolar disorder. *Psychiatry Clin Neurosci.* **71**, 77–103 (2016).
- Kauwe, J. S. K. et al. Genome-Wide Association Study of CSF Levels of 59 Alzheimer's Disease Candidate Proteins: Significant Associations with Proteins Involved in Amyloid Processing and Inflammation. *PLoS Genet.* **10**, e1004758 (2014).
- Dehghan, A. et al. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation* **123**, 731–738 (2011).
- Kjaergaard, A. D., Johansen, J. S., Nordestgaard, B. G. & Bojesen, S. E. Genetic variants in CH3L1 influencing YKL-40 levels: resequencing 900 individuals and genotyping 9000 individuals from the general population. *J. Med. Genet.* **50**, 831–837 (2013).
- Ryden, E. et al. A history of childhood attention-deficit hyperactivity disorder (ADHD) impacts clinical outcome in adult bipolar patients regardless of current ADHD. *Acta Psychiatr. Scandinavica* **120**, 239–246 (2009).
- Ryden, E., Johansson, C., Blennow, K. & Landén, M. Lower CSF HVA and 5-HIAA in bipolar disorder type 1 with a history of childhood ADHD. *J. Neural Transm. (Vienna, Austria: 1996)* **116**, 1667–1674 (2009).
- Jakobsson, J. et al. Altered concentrations of amyloid precursor protein metabolites in the cerebrospinal fluid of patients with bipolar disorder. *Neuropsychopharmacol.: Off. Publ. Am. Coll. Neuropsychopharmacol.* **38**, 664–672 (2013).
- Rolstad, S. et al. Cognitive performance and cerebrospinal fluid biomarkers of neurodegeneration: a study of patients with bipolar disorder and healthy controls. *PLoS ONE* **10**, e0127100 (2015).
- Abecasis, G. R. et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
- Song, J. et al. Genome-wide association study identifies SESTD1 as a novel risk gene for lithium-responsive bipolar disorder. *Mol. Psychiatry* **21**, 1290–1297 (2016).
- Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M. J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet. Epidemiol.* **32**, 381–385 (2008).
- Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).

36. UCSC Genome Browser. UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly [Internet]. http://genome-euro.ucsc.edu/cgi-bin/hgTracks?db=hg19&lastVirtModeType=default&lastVirtModeExtraState=&virtModeType=default&virtMode=0&nonVirtPosition=&position=chr2%3A125382755%2D125383255&hgid=222647055_OE4mND9W80rpAVkqZeEOh3xgK9 (2017).
37. Lee, J. J. et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
38. Yu, H. et al. Five novel loci associated with antipsychotic treatment response in patients with schizophrenia: a genome-wide association study. *Lancet Psychiatry* **5**, 327–338 (2018).
39. Djurovic, S. et al. A genome-wide association study of bipolar disorder in Norwegian individuals, followed by replication in Icelandic sample. *J. Affect. Disord.* **126**, 312–316 (2010).
40. Akiyama, M. et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nat. Genet.* **49**, 1458–1467 (2017).
41. Feitosa, M. F. et al. Novel genetic associations for blood pressure identified via gene-alcohol interaction in up to 570K individuals across multiple ancestries. *PLoS ONE* **13**, e0198166 (2018).
42. Nagel, M. et al. Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat. Genet.* **50**, 920–927 (2018).
43. Li, Z. et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat. Genet.* **49**, 1576–1583 (2017).
44. Winkler, T. W. et al. The influence of age and sex on genetic associations with adult body size and shape: a large-scale genome-wide interaction study. *PLOS Genet.* **11**, e1005378 (2015).
45. Armstrong, D. L. et al. GWAS identifies novel SLE susceptibility genes and explains the association of the HLA region. *Genes Immun.* **15**, 347–354 (2014).
46. National Center for Biotechnology Information. Gene [Internet]. <https://www.ncbi.nlm.nih.gov/gene/23314> (2017).
47. Walker, R. M. et al. DNA methylation in a Scottish family multiply affected by bipolar disorder and major depressive disorder. *Clin. Epigenetics* **8**, 5 (2016).
48. Zhao, W. et al. A new bioinformatic insight into the associated proteins in psychiatric disorders. *SpringerPlus* **5**, 1967 (2016).
49. Bergen, S. E. et al. Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. *Mol. Psychiatry* **17**, 880–886 (2012).
50. Etain, B. et al. Genome-wide scan for genes involved in bipolar affective disorder in 70 European families ascertained through a bipolar type I early-onset proband: supportive evidence for linkage at 3p14. *Mol. Psychiatry* **11**, 685–694 (2006).
51. Le-Niculescu, H. et al. Convergent functional genomics of genome-wide association data for bipolar disorder: Comprehensive identification of candidate genes, pathways and mechanisms. *Am. J. Med. Genet. Part B: Neuropsychiatr. Genet.* **150B**, 155–181 (2009).
52. Nurnberger, J. I. Jr. et al. Identification of pathways for bipolar disorder: a meta-analysis. *JAMA Psychiatry* **71**, 657–664 (2014).
53. Maes, H. H., Neale, M. C. & Eaves, L. J. Genetic and environmental factors in relative body weight and human adiposity. *Behav. Genet.* **27**, 325–351 (1997).
54. Allin, K. H., Bojesen, S. E. & Nordestgaard, B. G. Baseline C-reactive protein is associated with incident cancer and survival in patients with cancer. *J. Clin. Oncol.* **27**, 2217–2224 (2009).
55. Dehghan, A. et al. Genetic variation, C-reactive protein levels, and incidence of diabetes. *Diabetes* **56**, 872–878 (2007).
56. Sesso, H. D. et al. C-reactive protein and the risk of developing hypertension. *JAMA* **290**, 2945–2951 (2003).
57. Danesh, J. et al. C-reactive protein and other circulating markers of inflammation in the prediction of coronary heart disease. *N. Engl. J. Med.* **350**, 1387–1397 (2004).
58. Kaptoge, S. et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *Lancet (Lond., Engl.)* **375**, 132–140 (2010).
59. Naitza, S. et al. A genome-wide association scan on the levels of markers of inflammation in sardinians reveals associations that underpin its complex regulation. *PLoS Genet.* **8**, e1002480 (2012).
60. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat. Genet.* **42**, 441–447 (2010).
61. Treutlein, J. & Rietschel, M. Genome-wide association studies of alcohol dependence and substance use disorders. *Curr. Psychiatry Rep.* **13**, 147–155 (2011).