bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Genome-wide association study identifies 30 Loci Associated with Bipolar Disorder

Eli A Stahl1,2,3†&, Gerome Breen4,5†, Andreas J Forstner6,7,8,9,10†, Andrew McQuillin11†, Stephan Ripke12,13,14†, Vassily Trubetskoy13, Manuel Mattheisen15,16,17,18,19, Yunpeng Wang20,21, Jonathan R I Coleman4,5, Héléna A Gaspar4,5, Christiaan A de Leeuw22, Stacy Steinberg23, Jennifer M Whitehead Pavlides24, Maciej Trzaskowski25, Tune H Pers3,26, Peter A Holmans27, Liam Abbott12, Esben Agerbo19,28,29, Huda Akil30, Diego Albani31, Ney Alliey-Rodriguez32, Thomas D Als15,16,19, Adebayo Anjorin33, Verneri Antilla14, Swapnil Awasthi13, Judith A Badner34, Marie Bækvad-Hansen19,35, Jack D Barchas36, Nicholas Bass11, Michael Bauer37, Richard Belliveau12, Sarah E Bergen38, Carsten Bøcker Pedersen19,28,29, Erlend Bøen39, Marco Boks40, James Boocock41, Monika Budde42, William Bunney43, Margit Burmeister44, Jonas Bybjerg-Grauholm19,35, William Byerley45, Miquel Casas46,47,48,49, Felecia Cerrato12, Pablo Cervantes50, Kimberly Chambert12, Alexander W Charney2, Danfeng Chen12, Claire Churchhouse12,14, Toni-Kim Clarke51, William Coryell52, David W Craig53, Cristiana Cruceanu50,54, David Curtis55,56, Piotr M Czerski57, Anders M Dale58,59,60,61, Simone de Jong4,5, Franziska Degenhardt8,9, Jurgen Del-Favero62, J Raymond DePaulo63, Srdjan Djurovic64,65, Amanda L Dobbyn1,2, Ashley Dumont12, Torbjørn Elvsåshagen66,67, Valentina Escott-Price27, Chun Chieh Fan61, Sascha B Fischer6,10, Matthew Flickinger68, Tatiana M Foroud69, Liz Forty27, Josef Frank70, Christine Fraser27, Nelson B Freimer71, Louise Frisén72,73,74, Katrin Gade42,75, Diane Gage12, Julie Garnham76, Claudia Giambartolomei41, Marianne Giørtz Pedersen19,28,29, Jaqueline Goldstein12, Scott D Gordon77, Katherine Gordon-Smith78, Elaine K Green79, Melissa J Green80, Tiffany A Greenwood60, Jakob Grove15,16,19,81, Weihua Guan82, José Guzman Parra83, Marian L Hamshere27, Martin Hautzinger84, Urs Heilbronner42, Stefan Herms6,8,9,10, Maria Hipolito85, Per Hoffmann6,8,9,10, Dominic Holland58,86, Laura Huckins1,2, Stéphane Jamain87,88, Jessica S Johnson1,2, Anders Juréus38, Radhika Kandaswamy4, Robert Karlsson38, James L Kennedy89,90,91,92, Sarah Kittel-Schneider93, Sarah V Knott78, James A Knowles94,95, Manolis Kogevinas96, Anna C Koller8,9, Ralph Kupka97,98,99, Catharina Lavebratt72, Jacob Lawrence100, William B Lawson85, Markus Leber101, Phil H Lee12,14,102, Shawn E Levy103, Jun Z Li104, Chunyu Liu105, Susanne Lucae106, Anna Maaser8,9, Donald J MacIntyre107,108, Pamela B Mahon63,109, Wolfgang Maier110, Lina Martinsson73, Steve McCarroll12,111, Peter McGuffin4, Melvin G McInnis112, James D McKay113, Helena Medeiros95, Sarah E Medland77, Fan Meng30,112, Lili Milani114, Grant W Montgomery25, Derek W Morris115,116, Thomas W Mühleisen6,117, Niamh Mullins4, Hoang Nguyen1,2, Caroline M Nievergelt60,118, Annelie Nordin Adolfsson119, Evaristus A Nwulia85, Claire O'Donovan76, Loes M Olde Loohuis71, Anil P S Ori71, Lilijana Oruc120, Urban Ösby121, Roy H Perlis122,123, Amy Perry78, Andrea Pfennig37, James B Potash63, Shaun M Purcell2,109, Eline J Regeer124, Andreas Reif93, Céline S Reinbold6,10, John P Rice125, Fabio Rivas83, Margarita Rivera4,126, Panos Roussos1,2,127, Douglas M Ruderfer128, Euijung Ryu129, Cristina Sánchez-Mora46,47,49, Alan F Schatzberg130, William A Scheftner131, Nicholas J Schork132, Cynthia Shannon Weickert80,133, Tatyana Shehktman60, Paul D Shilling60, Engilbert Sigurdsson134, Claire Slaney76, Olav B Smeland58,135,136, Janet L Sobell137, Christine Søholm Hansen19,35, Anne T Spijker138, David St Clair139, Michael Steffens140, John S Strauss91,141, Fabian Streit70, Jana Strohmaier70, Szabolcs Szelinger142, Robert C Thompson112, Thorgeir E Thorgeirsson23, Jens Treutlein70, Helmut Vedder143, Weiqing Wang1,2, Stanley J Watson112, Thomas W Weickert80,133, Stephanie H Witt70, Simon Xi144, Wei Xu145,146, Allan H Young147, Peter Zandi148, Peng Zhang149, Sebastian Zollner112, Rolf Adolfsson119, Ingrid Agartz17,39,150, Martin Alda76,151, Lena Backlund73, Bernhard T Baune152, Frank Bellivier153,154,155,156, Wade H Berrettini157, Joanna M Biernacka129, Douglas H R Blackwood51, Michael Boehnke68, Anders D Børglum15,16,19, Aiden Corvin116, Nicholas Craddock27, Mark J Daly12,14, Udo Dannlowski158, Tõnu Esko3,111,114,159, Bruno Etain153,155,156,160, Mark Frye161, Janice M Fullerton133,162, Elliot S Gershon32,163, Michael Gill116, Fernando Goes63, Maria Grigoroiu-Serbanescu164, Joanna Hauser57, David M Hougaard19,35, Christina M Hultman38, Ian Jones27, Lisa A Jones78, René S Kahn2,40, George Kirov27, Mikael Landén38,165, Marion Leboyer88,153,166, Cathryn M Lewis4,5,167, Qingqin S Li168, Jolanta Lissowska169, Nicholas G Martin77,170, Fermin Mayoral83, Susan L McElroy171, Andrew M McIntosh51,172, Francis J McMahon173, Ingrid Melle174,175, Andres Metspalu114,176, Philip B Mitchell80, Gunnar Morken177,178, Ole Mors19,179, Preben Bo Mortensen15,19,28,29, Bertram Müller-Myhsok54,180,181, Richard M Myers103, Benjamin M Neale3,12,14, Vishwajit Nimgaonkar182, Merete Nordentoft19,183, Markus M Nöthen8,9, Michael C O'Donovan27, Ketil J Oedegaard184,185, Michael J Owen27, Sara A Paciga186, Carlos Pato95,187, Michele T Pato95, Danielle Posthuma22,188, Josep Antoni Ramos-Quiroga46,47,48,49, Marta Ribasés46,47,49, Marcella Rietschel70, Guy A Rouleau189,190, Martin Schalling72, Peter R Schofield133,162, Thomas G Schulze42,63,70,75,173, Alessandro Serretti191, Jordan W Smoller12,192,193, Hreinn Stefansson23, Kari Stefansson23,194, Eystein Stordal195,196, Patrick F Sullivan38,197,198, Gustavo Turecki199, Arne E Vaaler200, Eduard Vieta201, John B Vincent141, Thomas Werge19,202,203, John I Nurnberger204, Naomi R Wray24,25, Arianna Di Florio27,198, Howard J Edenberg205, Sven Cichon6,8,10,117, Roel A Ophoff40,41,71, Laura J Scott68, Ole A Andreassen135,136, John Kelsoe60*&, Pamela Sklar1,2*^

† Equal contribution * Co-last authors
& Correspondence to: jkelsoe@ucsd.edu or eli.stahl@mssm.edu
^ deceased. This paper is dedicated to the memory of Psychiatric Genomics Consortium founding member and Bipolar disorder working group co-chair Pamela Sklar

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

ABSTRACT:

Bipolar disorder is a highly heritable psychiatric disorder that features episodes of mania and depression. We performed the largest genome-wide association study to date, including 20,352 cases and 31,358 controls of European descent, with follow-up analysis of 822 sentinel variants at loci with $P<1x10^{-4}$ in an independent sample of 9,412 cases and 137,760 controls. In the combined analysis, 30 loci reached genome-wide significant evidence for association, of which 20 were novel. These significant loci contain genes encoding ion channels and neurotransmitter transporters (*CACNA1C*, *GRIN2A*, *SCN2A, SLC4A1*), synaptic components (*RIMS1, ANK3*), immune and energy metabolism components. Bipolar disorder type I (depressive and manic episodes; ~73% of our cases) is strongly genetically correlated with schizophrenia whereas bipolar disorder type II (depressive and hypomanic episodes; ~17% of our cases) is more strongly correlated with major depressive disorder. These findings address key clinical questions and provide potential new biological mechanisms for bipolar disorder.

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

INTRODUCTION

Bipolar disorder (BD) is a severe neuropsychiatric disorder characterized by recurrent episodes of mania and depression which affect thought, perception, emotion, and social behaviour. A lifetime prevalence of 1-2%, elevated morbidity and mortality, onset in young adulthood, and a frequently chronic course make BD a major public health problem and a leading cause of the global burden of disease [1]. Clinical, twin and molecular genetic data all strongly suggest that BD is a multifactorial disorder [2]. Based on twin studies, the overall heritability of BD has been estimated to be more than 70% [3,4], suggesting a substantial involvement of genetic factors in the development of the disorder, although non-genetic factors also influence risk.

BD can be divided into two main clinical subtypes [5,6]: bipolar I disorder (BD1) and bipolar II disorder (BD2). In BD1, manic episodes typically alternate with depressive episodes during the course of illness. Diagnosis of BD2 is based on the lifetime occurrence of at least one depressive and one hypomanic (but no manic) episode. Although modern diagnostic systems retain the Kraepelinian dichotomy [7] between BD and schizophrenia, the distinction between the two disorders is not always clear-cut, and patients who display clinical features of both disorders may receive a diagnosis of schizoaffective disorder (SAB). Likewise, in genetic studies the two diagnoses are usually treated separately, although recent epidemiological and molecular genetic studies provide strong evidence for some overlap between the genetic contributions to their etiology [2,8].

Recent genome-wide association studies (GWAS) in BD have identified a number of significant associations between disease status and common genetic variants [9–23]. The first large collaborative BD GWAS by the multinational Psychiatric Genomics Consortium (PGC) Bipolar Disorder Working Group comprised 7,481 BD patients and 9,250 controls and identified four genome-wide significant loci [9]. Three subsequent meta-analyses that included the PGC BD data

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

[10,12,18] identified an additional 5 loci.

Estimates of the proportion of variance in liability attributable to common variants genome-wide (SNP-heritability) indicate that ~30% of the heritability for BD is due to common genetic variants [8]. To date, only a small fraction of this heritability is explained by associated loci, but results from other human complex traits suggest that many more will be identified by increasing the sample size of GWAS [24]. Here, we report the second GWAS of the PGC Bipolar Disorder Working Group, comprising 20,352 cases and 31,358 controls of European descent in a single, systematic analysis, with follow up of top findings in an independent sample of 9,412 cases and 137,760 controls. Some of our findings reinforce specific hypotheses regarding BD neurobiology; however, the majority of the findings suggest new biological insights.

RESULTS

GWAS of bipolar disorder (BD)

We performed a GWAS meta-analysis of 32 cohorts from 14 countries in Europe, North America and Australia (**Supplementary Table 1A**), totaling 20,352 cases and 31,358 controls of European descent (effective sample size 46,582). This is the largest GWAS of BD to date and includes 6,328 case and 7,963 control samples not previously reported, a 2.7-fold increase in the number of cases compared to our previous GWAS [9]. We imputed variant dosages using the 1,000 Genomes reference panel (see Methods), retaining association results for 9,372,253 autosomal variants with imputation quality score INFO > 0.3 and minor allele frequency ≥ 1% in both cases and controls. We performed logistic regression of case status on imputed variant dosage using genetic ancestry covariates. The resulting genomic inflation factor (λ_{GC}) was 1.23 and scaled to 1,000 cases and 1,000 controls (λ_{1000}) was 1.01 (**Supplementary Figure 1**). The LD-score regression intercept did not significantly differ from one, indicating that the observed genomic

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

inflation is indicative of polygenicity rather than stratification or cryptic population structure [25]. The LD-score regression SNP-heritability estimates for BD were 0.17-0.23 (on the liability scale, assuming population lifetime risk of 0.5-2%). See **Supplementary Table 1A**, **Online Methods** and **Supplementary Note** for sample and method details.

We find a marked increase in phenotypic variance explained by genomewide polygenic risk scores (PRS) compared to previous publications (sample size weighted mean observed Nagelkerke's R^2 = 0.08 across datasets, liability scale R^2=0.04, for P-threshold ≤ 0.01; **Supplementary Figure 2** and **Supplementary Table 2**). Among the different datasets, we observed no association between the PRS and: (i) the gender distribution of the BD cases (p=0.51); (ii) the proportion of cases with psychosis (p=0.61); (iii) the proportion with a family history of BD (p=0.82); or (iv) the median age of onset for BD (p=0.64). In our primary genome-wide analysis, we identified 19 loci exceeding genome-wide significance ($P< 5x10^{-8}$).

Follow-up of suggestive loci in additional samples

We meta-analyzed lead variants that were significant at $P<1x10^{-4}$ in our discovery meta-analysis, (a total of 794 autosomal and 28 X chromosome variants) with follow-up samples totaling 9,412 cases and 137,760 controls of European ancestry (**Supplementary Note** and **Supplementary Table 1B**). Thirty autosomal loci achieved combined sample genome-wide significance ($P< 5x10^{-8}$) (**Figure 1**, **Table 1**, **Supplementary Figure 3**, **Supplementary Table 3**). These include 19 loci that were significant only in the combined analysis, of which three were reported to have genome-wide significant SNPs in previous studies (*ADCY2* [18], *POU3F2* [18], *ANK3* [12,18]), and 11 that were significant in our GWAS. Eight variants were genome-wide significant in the GWAS but not in the combined analysis. Using effect sizes corrected for winner's curse [26,27] for each of the 19 variants with GWAS $P<5x10^{-8}$, we found that 11 variants achieving genome-wide significance in

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

our combined analysis is within the expected range (Poisson binomial test P = 0.29, **Supplementary Note** and **Supplementary Figure 4**).

Lead variants for the 30 loci achieving genome-wide significance in the combined analysis are shown in **Table 1A**. We show results in **Table 1B** for 8 additional loci with $P < 5x10^{-8}$ in our discovery GWAS but not in the combined analysis. Results for all variants tested in the follow-up study are presented in **Supplementary Table 3**. We refer to loci by the gene name attributed in previous BD GWAS publications, or by the name of the closest gene for novel loci, without implication that the named gene is causal. Of the 30 genome-wide significant loci from our combined analysis, 20 are novel BD risk loci. In **Supplementary Table 4**, we present detailed descriptions of the associated loci and genes, with bioinformatic and literature evidence for their potential roles in BD.

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

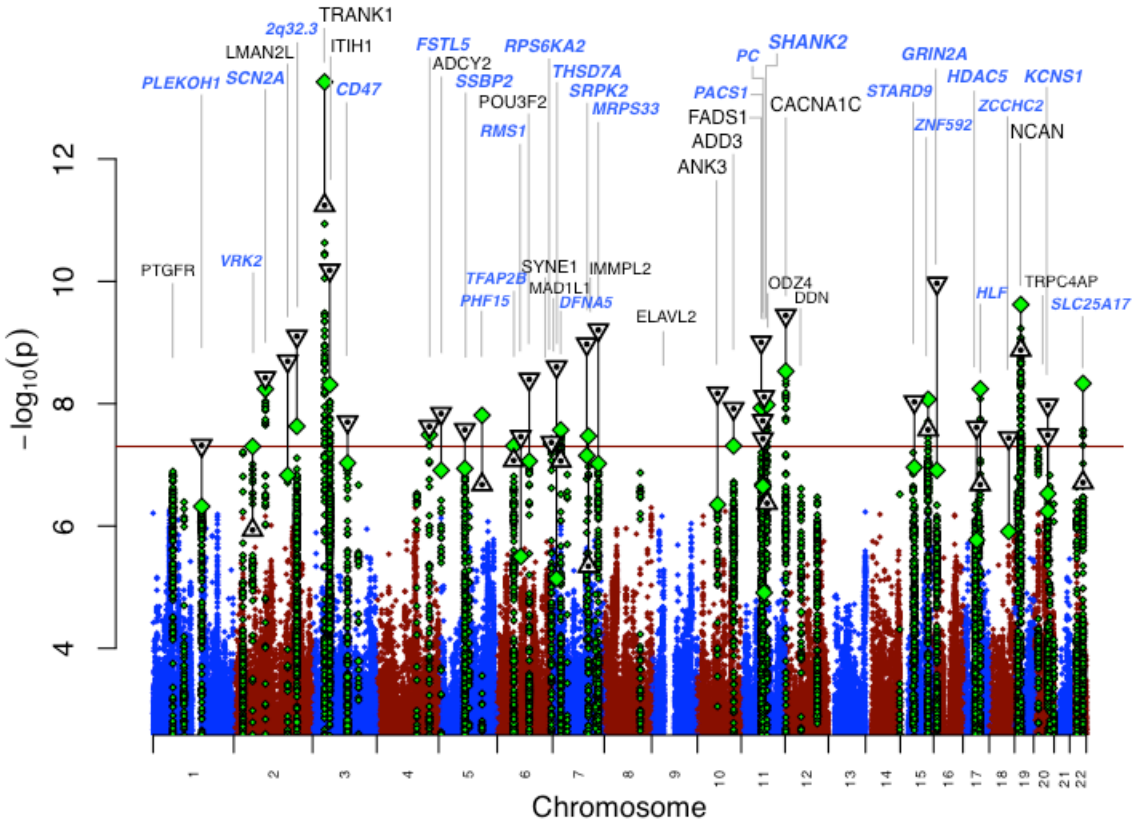

Figure 1. Manhattan plot for our primary genomewide association analysis of 20,352 cases and 31,358 controls. GWAS -$\log_{10}$P-values are plotted for all SNPs across chromosomes 1-22 (diamonds, green for loci with lead SNP GWAS P < 10^{-6}). Combined GWAS+followup -$\log_{10}$P-values for lead SNPs reaching genome-wide significance in either GWAS or combined analysis (triangles, inverted if GWAS+followup -$\log_{10}$P > GWAS -$\log_{10}$P). Labels correspond to gene symbols previously reported for published loci (black) and the nearest genes for novel loci (blue), at top if GWAS+followup P < 5×10^{-8}.

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Table 1. Genome-wide significant bipolar disorder risk loci

					GWAS Meta-analysis			Follow-up samples		Combined	
Locus Name*	Lead SNP	CHR	BP	A1/A2	Freq. A1	OR	P-value	OR	P-value	OR	P-value
A. Thirty loci with lead SNP P < 5x10-8 in combined GWAS+followup analysis											
1,PLEKHO1	rs7544145	1	150,138,699	T/C	0.81	1.095	4.8E-07	1.064	0.021	1.085	**4.8E-08**
2,LMAN2L**	chr2_97376407_I	2	97,376,407	I/D	0.34	0.92	**5.8E-09**	0.96	0.059	0.93	**3.8E-09**
3,SCN2A	rs17183814	2	166,152,389	A/G	0.075	0.87	1.5E-07	0.89	0.0033	0.88	**2.0E-09**
4,[Intergenic]***	chr2_194465711_D	2	194,465,711	I/D	0.41	0.93	**2.3E-08**	0.95	0.0063	0.93	**7.9E-10**
5,TRANK1**	rs9834970	3	36,856,030	T/C	0.51	0.90	**5.5E-14**	0.98	0.30	0.93	**5.7E-12**
6,ITIH1**	rs2302417	3	52,814,256	A/T	0.49	0.92	**4.9E-09**	0.94	0.0024	0.93	**6.6E-11**
7,CD47	rs3804640	3	107,793,709	A/G	0.53	1.075	9.3E-08	1.044	0.032	1.065	**2.0E-08**
8,FSTL5	rs11724116	4	162,294,038	T/C	0.16	0.90	**3.3E-08**	0.95	0.061	0.92	**2.4E-08**
9,ADCY2**	chr5_7587236_D	5	7,587,236	I/D	0.82	0.91	1.2E-07	0.94	0.023	0.92	**1.5E-08**
10,SSBP2	rs10035291	5	80,796,368	T/C	0.68	1.081	1.1E-07	1.047	0.036	1.070	**2.7E-08**
11,RIMS1	chr6_72519394_D	6	72,519,394	D/I	0.44	1.066	3.1E-06	1.062	0.0033	1.064	**3.5E-08**
12,POU3F2**	rs2388334	6	98,591,622	A/G	0.52	0.93	8.6E-08	0.95	0.010	0.94	**4.0E-09**
13,RPS6KA2	rs10455979	6	166,995,260	C/G	0.53	0.93	**4.6E-08**	0.97	0.092	0.94	**4.3E-08**
14,THSD7A	rs113779084	7	11,871,787	A/G	0.30	1.068	7.3E-06	1.095	5.7E-05	1.076	**2.5E-09**
15,SRPK2	rs73188321	7	105,048,158	T/C	0.33	0.92	7.0E-08	0.94	0.0030	0.92	**1.1E-09**
16,MRPS33	chr7_140700006_I	7	140,700,006	D/I	0.25	0.92	9.4E-08	0.93	0.0015	0.92	**6.2E-10**
17,ANK3**	rs10994318	10	62,125,856	C/G	0.057	1.151	4.5E-07	1.130	0.0041	1.145	**6.8E-09**
18,ADD3**	chr10_111745562_I	10	111,745,562	I/D	0.16	1.105	**5.0E-08**	1.059	0.034	1.090	**1.2E-08**
19,FADS2**	rs12226877	11	61,591,907	A/G	0.29	1.095	**1.2E-08**	1.062	0.015	1.085	**9.9E-10**
20,PACS1	rs10896090	11	65,945,186	A/G	0.81	1.094	2.1E-07	1.062	0.018	1.084	**1.9E-08**
21,PC	rs7122539	11	66,662,731	A/G	0.35	0.93	2.2E-07	0.96	0.030	0.94	**3.8E-08**
22,SHANK2	rs12575685	11	70,517,927	A/G	0.31	1.066	1.2E-05	1.088	1.1E-04	1.073	**7.7E-09**
23,CACNA1C**	rs10744560	12	2,387,099	T/C	0.34	1.087	**2.9E-09**	1.052	0.017	1.076	**3.6E-10**
24,STARD9	rs4447398	15	42,904,904	A/C	0.12	1.112	1.1E-07	1.072	0.016	1.099	**9.4E-09**
25,ALPK3	chr15_85357857_I	15	85,357,857	I/D	0.28	0.92	**8.5E-09**	0.97	0.16	0.93	**2.7E-08**
26GRIN2A	rs11647445	16	9,926,966	T/G	0.65	0.93	1.2E-07	0.93	1.96E-04	0.93	**1.1E-10**
27,HDAC5	rs112114764	17	42,201,041	T/G	0.69	0.93	1.7E-06	0.94	0.0042	0.93	**2.5E-08**
28,ZCCHC2	rs11557713	18	60,243,876	A/G	0.29	1.074	1.2E-06	1.059	0.0077	1.069	**3.6E-08**
29,NCAN**	rs111444407	19	19,358,207	T/C	0.15	1.124	**2.4E-10**	1.040	0.15	1.097	**1.3E-09**
30,STK4	chr20_43682549_I	20	43,682,549	I/D	0.28	0.923	3.0E-07	0.942	0.009	0.929	**1.1E-08**
B. Additional loci with lead SNP P < 5x10-8 in GWAS analysis											
TFAP2B	rs55648125	6	50816718	A/G	0.90	0.89	**4.9E-08**	0.95	0.14	0.91	8.5E-08
DFNA5	rs17150022	7	24771777	T/C	0.88	0.89	**2.7E-08**	0.96	0.17	0.91	8.6E-08
SLC25A17	rs138321	22	41209304	A/G	0.50	1.083	**4.7E-09**	1.012	0.55	1.060	1.9E-07
HLF	rs884301	17	53367464	T/C	0.37	1.084	**5.8E-09**	1.013	0.52	1.061	2.1E-07
PHF15	rs329319	5	133906609	A/G	0.43	1.082	**1.5E-08**	1.019	0.36	1.061	2.1E-07
ODZ4**	rs73496688	11	79156748	A/T	0.14	1.11	**1.0E-08**	1.016	0.58	1.083	4.2E-07
[Intergenic]***	rs57681866	2	57975714	A/G	0.06	0.85	**5.0E-08**	0.97	0.45	0.89	1.2E-06
[Intergenic]***	rs13231398	7	110197412	C/G	0.11	0.89	**3.4E-08**	0.998	0.95	0.92	4.6E-06

* Loci are numbered 1 to 30, ordered by genomic position, with previously reported gene name for published loci

** Previously published and named loci. (Locus 12 would be named as Intergenic, nearest gene is POU3F2 691Kb.)

*** Intergenic loci nearest genes: Locus 4 PCGEM1 824kb, Table 1B chr2 locus VRK2 298Kb, Table 1B chr7 IMMP2L 106Kb.

We next asked if the variants tested in the follow-up samples were, in aggregate, consistent with the presence of additional sub genome-wide significant BD association signals. After excluding 47 variants that were genome-wide significant in our GWAS, our combined analysis or previous BD GWAS, 775 variants remained in our follow-up experiment. 551 variants had the same direction of effect in the discovery GWAS and follow-up samples (71% compared to a null expectation of 50%, sign test $P < 2.2 \times 10^{-16}$), and 110 variants had the same direction of effect and were nominally significant ($p<0.05$) in the follow-up samples (14% compared to an expected value of 2.5% , binomial test $P < 2.2 \times 10^{-16}$). This consistency between our GWAS and follow-up samples suggests that many true BD associations exist among these variants.

To identify additional independent signals, we conducted conditional analyses across each of the 30 significant BD loci (**Supplementary Table 5**). We used the effective number of

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

independent variants based on LD structure within loci [28] to calculate a multiple test-corrected significance threshold ($P=1.01\times10^{-5}$, see **Supplementary Note**). One locus showed evidence for an independent association signal (rs114534140 in locus #8, *FSTL5*; $P_{conditional} = 2\times10^{-6}$). At one locus (#30, STK4 on chr 20), we found two SNPs with genome-wide significance in low LD ($r^2 < 0.1$); however, conditional analysis showed that their associations were not independent. Thus only the *FSTL5* locus demonstrated clear evidence of more than one independent association.

Shared loci and genetic correlations with schizophrenia, depression and other GWAS traits

We next examined the genetic relationships of BD to other psychiatric disorders and traits. Of the 30 genome-wide significant BD loci, 8 also harbor schizophrenia (SCZ) associations [29–31]. Based on conditional analyses the BD and SCZ associations appear to be independent at 3 of the 8 shared loci (*NCAN*, *TRANK1* and chr7q22.3:105Mb loci) (**Supplementary Table 6**). No genome-wide significant BD locus overlapped with those identified for major depression (DEPR), including 44 risk loci identified in the most recent PGC study based on 130,664 depression cases and 330,470 controls[32], and those reported in a large study of depressive symptoms or subjective well-being [33]. As previously reported [34], we found substantial and highly significant genetic correlations between BD and SCZ (LD-score regression estimated genetic correlation $r_g = 0.70$, se = 0.020) and between BD and DEPR ($r_g = 0.35$, se = 0.026) The BD and DEPR genetic correlation was similar to that observed for SCZ and DEPR ($r_g = 0.34$, se = 0.025) (**Supplementary Table 7A**).

We found significant genetic correlations between BD and other psychiatric-relevant traits (**Supplementary Table 7B**), including with autism spectrum disorder [8] ($r_g = 0.18$, $P=2\times10^{-4}$), anorexia nervosa [35] ($r_g = 0.23$, $P=9\times10^{-8}$), and subjective well-being [33] ($r_g = -0.22$, $P=4\times10^{-7}$). There was suggestive positive overlap with anxiety disorders ($r_g=0.21$, P=0.04) [36] and neuroticism

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

(r_g=0.12, P=0.002) [37]. Significant r_gs were seen with measures of education: college attendance [38] (r_g = 0.21, P=1=x10^{-7}) and education years [39] (r_g=0.20, P=6x10^{-14}), but not with childhood IQ [40] (r_g=0.05, P=0.5) or intelligence [41] (r_g=-0.05, P=0.08). Among a large number of BD risk locus SNPs associated with additional traits from GWAS catalog, we found a handful of loci with non-independent associations (in one overlapping locus each with educational attainment, biliary atresia, bone mineral density, lipid-related biomarkers) (**Supplementary Table 6**). Biliary atresia and lipid- related biomarkers, however, did not show significant genetic correlation with BD (**Supplementary Table 7B**).

BD subtype GWAS

We performed secondary GWAS focusing on three clinically recognized subtypes of bipolar disorder: BD1 (n=14,879 cases), BD2 (n=3,421 cases), and SAB (n=977 cases) (**Supplementary Note**, **Supplementary Tables 1A** & **8, Supplementary Figure 5**). We observed variants in 14 loci with genome-wide significance for BD1, 10 of which were in genome-wide significant loci in the combined BD GWAS analysis. Not surprisingly given the sample overlap, 3 of the 4 remaining loci genome-wide significant for BD1 have P < 10^{-6} in either our GWAS or combined analysis. The remaining locus (MAD1L1, chr7:1.9Mb, GWAS P = 2.4x10^{-6}) was recently published in two BD GWAS that included Asian samples [42,43]. We did not observe genome-wide significant results for the smaller BD2 and SAB analyses. BD1, BD2 and SAB all have significant common variant heritabilities (BD1 h^2_{snp} = 0.25, se = 0.014, P = 3.2x10^{-77}; BD2 h^2_{snp} = 0.11, se = 0.028, P = 5.8x10^{-5}; SAB h^2_{snp} = 0.25, se = 0.10, P = 0.0071). Genetic correlations among BD subtypes show that these represent closely related, yet partially distinct, phenotypes (**Supplementary Table 9)**.

Polygenic risk scores and genetic correlations provide support for a continuum of SCZ-BD1-BD2-DEPR genetic effects, with significantly greater genetic SCZ polygenic risk scores (PRS)

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

in BD1 cases than in BD2 cases (min P=5.6x10^{-17}, P threshold = 0.1), and greater DEPR PRS in BD2 cases than in BD1 cases (min P=8.5x10^{-10}, P threshold = 0.01) (**Figure 2**, **Supplementary Table 10**). Genetic correlations from LD-score regression support these results; genetic correlations were greater for SCZ with BD1 (r_g = 0.71, se = 0.025) than with BD2 (r_g = 0.51, se = 0.072), with P_{diff} = 0.0056, and were greater for DEPR with BD2 (r_g = 0.69, se = 0.093) than with BD1 (r_g = 0.30, se = 0.028), with P_{diff} = 2.9x10^{-5} (**Supplementary Table 9**).

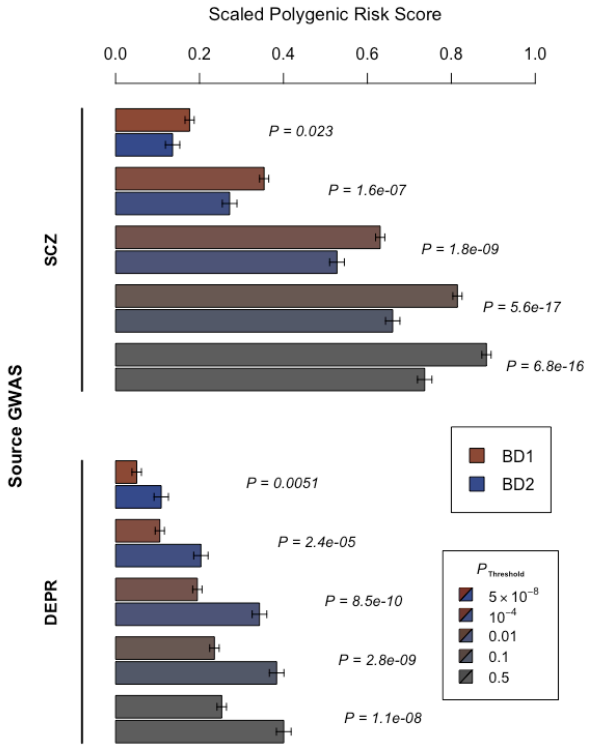

Figure 2. Association of BD1 and BD2 subtypes with schizophrenia (SCZ) and major depression (DEPR) polygenic risk scores (PRS). Shown are mean PRS values (1 s.e. error bars), adjusted for study and ancestry covariates and scaled to the PRS mean and sd in control subjects, in BD1 (red) and BD2 (blue) cases, for increasing source GWAS P-value thresholds (increasing grey) as indicated. P-values (italics) test BD1 vs BD2 mean PRS, in logistic regression of case subtype on PRS with covariates. Results are detailed in Supplementary Table 10.

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Systems biology and *in silico* functional analyses of BD GWAS results

To identify genes with functional variation in gene expression that might explain the associations, we used summary Mendelian randomization (SMR) [44] to integrate our BD discovery GWAS with eQTL data from brain dorsolateral prefrontal cortex [45] as well as a large-sample whole blood eQTL dataset [46] (**Supplemental Table 11**). SMR identified six transcriptome-wide significant genes without signs of heterogeneity between GWAS and eQTL association signals. Among these, four genes were present in four different loci from our combined BD GWAS and follow-up sample meta-analysis: *LMAN2L* (blood), *FADS1* (brain), *NMB* (blood) and *C17ORF65* (blood).

We tested for functional genomic enrichment in our BD GWAS using partitioned LD-score regression [47] (**Supplementary Note**, **Supplementary Table 12**). Annotations tested included open chromatin DHS peaks in a range of tissues [48], genic annotations, conservation, and a number of functional genomic annotations across tissues. SNP-based BD heritability was most substantially enriched in open chromatin annotations in central nervous system (proportion SNPs = 0.14, proportion h^2_{snp} = 0.60, enrichment =3.8, P = 4.2 x 10^{-17}) . We also used DEPICT [49] to test for expression of BD associated genes across tissues, and found significant enrichment of central nervous system (P <= 1.3x10^{-3}, FDR < 0.01) and neurosecretory system (P <= 2.0x10^{-6}, FDR < 0.01) genes (**Supplementary Table 13**).

Finally, we used MAGMA [50] to conduct a gene-wise BD GWAS and to test for enrichment of pathways curated from multiple sources (see **Supplementary Note**). We note that significance levels were assigned to genes by physical proximity of SNPs, and do not imply that significant genes are causal for BD. Genic association results included 154 Bonferroni significant genes (MAGMA P_JOINT < 2.8x10^{-6}), including 82 genes in 20 genome-wide significant loci, and 73 genes in 27 additional loci that did not reach genome-wide significance in either our GWAS or

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

combined analysis (**Supplementary Table 14**). Nine related pathways were significantly enriched for genes with stronger BD associations ($P < 7.0x10^{-5}$, FDR < 0.05), including abnormal motor coordination/balance pathways (from mice), regulation of insulin secretion and endocannabinoid signaling pathways (**Supplementary Table 15, Supplementary Figure 6**).

DISCUSSION

We carried out the largest bipolar disorder (BD) GWAS to date and identified 30 genome-wide significant loci, including 20 novel BD risk loci. Previous BD GWAS have reported a total of 20 loci significantly associated with BD[9–23] ; twelve of these previously reported loci were not genome-wide significant in our GWAS meta analysis but had $P_{GWAS} \leq 1.3x10^{-5}$. Of the 19 loci identified in our discovery GWAS, only 11 were genome-wide significant in meta-analysis of our GWAS and follow-up samples. Although these results are not unexpected given small effect sizes and the winner's curse [27,51] (**Supplementary Note** and **Supplementary Figure 4**), genetic heterogeneity has been shown between BD GWAS cohorts[8]. We observed variable polygenic effects between BD subtypes and between cohorts in our study (**Figure 2**, **Supplementary Figure 2**, **Supplementary Tables 2** & **10**) and acknowledge a diversity of clinical case phenotypic criteria among cohorts in our study (**Supplementary Note**). Remarkably, our strongest association signal, observed at the *TRANK1* locus (rs9834970; $P_{combined}$ = 5.7E-12, OR = 0.93), exhibited significant heterogeneity among discovery GWAS cohorts (Cochran's Q $P = 1.9x10^{-4}$, and did not replicate in the follow-up sample (1-tailed $P_{followup}$ = 0.3) (**Supplementary Figure 3B & 3C**, fifth and first plots respectively). This locus has been observed in recent [11,12,17,18] but not earlier BD GWAS [9,13,20], surprisingly given its relatively large apparent effect size. Thus, complex polygenic architecture as well as phenotypic heterogeneity among BD GWAS cohorts may contribute to the inconsistency of genome-wide significant findings within and across BD GWAS studies. The

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

observed heterogeneity is a major challenge for GWAS of psychiatric disorders and calls for careful and systematic clinical assessment of cases and controls in addition to continued efforts to collect larger sample sizes.

Of the 30 BD associated loci, 8 also harbor associations [29–31] with schizophrenia (SCZ); however, conditional analyses suggest that the BD and SCZ associations at 3 of the 8 shared loci (in the *NCAN*, *TRANK1* and chr7q22.3 [105Mb] loci) may be independent (**Supplementary Table 6**). Differential BD and SCZ associations may represent opportunities to understand the genetic distinctions between these closely related and sometimes clinically difficult to distinguish disorders. We did not find BD loci that overlap with those associated with major depression[32].

The confirmed association within loci containing *CACNA1C* and other voltage-gated calcium channels supports the rekindled interest in calcium channel antagonists as potential treatments for BD with similar examination ongoing for other genes implicated by current GWAS [52]. These processes are important in neuronal hyperexcitability[53], an excess of which has been reported in iPSC derived neurons from BD patients, and which has been shown to be affected by the classic mood stabilizing drug lithium[54]. Other genes within novel associated loci include those coding for neurotransmitter channels (*GRIN2A*), ion channels and transporters (*SCN2A, SLC4A1*) and synaptic components (*RIMS1, ANK3*). Further study will confirm whether or not these are the causal genes in these loci.

The estimated variance explained by polygenic risk scores (PRS) based on our BD GWAS data is ~8% (observed scale; 4% on the liability scale [55]), an increase from 2.8% from our previous study [9]. Using PRS, we found that BD1 cases have significantly greater schizophrenia genetic risk than BD2 cases, while BD2 cases have significantly greater major depression genetic risk than BD1 cases, consistent with a spectrum of related psychiatric diagnoses[7,56]. We observe significant positive genetic correlations with educational attainment, but not with either adult or

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

childhood IQ, suggesting that the role of BD genetics in increased educational attainment may be independent of general intelligence. This result is inconsistent with suggestions from epidemiological studies [57], but in agreement with a recent clinical study [58].

In summary, findings from the largest genome-wide analysis of BD reveal an extensive polygenic genetic architecture of the disease, implicate brain calcium channels and neurotransmitter function in BD etiology, and confirm that BD is part of a spectrum of highly correlated psychiatric and mood disorders.

ONLINE METHODS

Methods

GWAS and follow-up cohorts. Our discovery GWAS sample was comprised of 32 cohorts from 14 countries in Europe, North America and Australia (**Supplementary Table 1A**), totaling 20,352 cases and 31,358 controls of European descent. A selected set of variants (see below) were tested in 7 follow-up cohorts of European descent (**Supplementary Table 1B**), totalling 9,025 cases and 142,824 controls (N_{eff} = 23,991). The **Supplementary Note** summarizes the source and inclusion/exclusion criteria for cases and controls for each cohort. All cohorts in the initial PGC BD paper were included [9]. Cases were required to meet international consensus criteria (DSM-IV, ICD-9, or ICD-10) for a lifetime diagnosis of BD established using structured diagnostic instruments from assessments by trained interviewers, clinician-administered checklists, or medical record review. In most cohorts, controls were screened for the absence of lifetime psychiatric disorders and randomly selected from the population.

GWAS cohort analysis We tested 20 principal components for association with BD using logistic regression; seven were significantly associated with phenotype and used in GWAS association

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

analysis (PCs 1-6, 19). In each cohort, we performed logistic regression association tests for BD with imputed marker dosages including 7 principal components to control for population stratification. For all GWAS cohorts, X-chromosome association analyses were conducted separately by sex, and then meta-analyzed across sexes. We also conducted BD1, BD2, and SAB GWAS, retaining only cohorts with at least 35 subtype cases and filtering SNPs for MAF > 0.02**.** Results were combined across cohorts using an inverse variance-weighted fixed effects meta-analysis [59]. We used Plink 'clumping' [60,61] to identify an LD-pruned set of discovery GWAS meta-analysis BD-associated variants ($P < 0.0001$, and distance >500kb or LD $r^2 < 0.1$, n variants =822) for analysis in the follow-up cohorts. Conditional analyses were conducted within each GWAS cohort and meta-analyzed as above.

Follow-up cohort analysis. In each follow-up cohort we performed BD association analysis of the 822 selected GWAS variants (when available) including genetic ancestry covariates, following QC and analysis methods of the individual study contributors. We performed inverse variance-weighted fixed-effects meta-analyses of the association results from the follow-up cohorts, and of the discovery GWAS and follow-up analyses.

Polygenic risk score (PRS) analyses. We tested PRS for our primary GWAS on each GWAS cohort as a target set, using a GWAS where the target cohort was left out of the meta-analysis (**Supplementary Table 2**). To test genetic overlaps with other psychiatric diseases, we calculated PRS for DEPR and SCZ in our GWAS cohort BD cases [62]. In pairwise case subtype analyses (**Figure 2**, **Supplementary Table 10**), we regressed subtype case status (BD1 n=8044, BD2 n=3,365, SAB n=977) on the PRS adjusting for ancestry principal components and a cohort indicator using logistic regression, and visualized covariate-adjusted PRS in BD1 and BD2 subtypes (**Figure 2**).

Linkage disequilibrium (LD) score regression. LD score regression [25,63] was used to conduct SNP-heritability analyses from GWAS summary statistics. LD score regression bivariate genetic

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

correlations attributable to genome-wide common variants were estimated between the full BD GWAS, BD subtype GWASs, and other traits and disorders with LD-Hub [63]. We also used LD score regression to partition heritability by genomic features [47].

Relation of BD GWA findings to tissue and cellular gene expression. We used partitioned LD score regression to evaluate which somatic tissues and brain tissues were enriched for BD heritability. [64] We used summary-data-based Mendelian randomization (SMR) [44] to identify loci with strong evidence of causality via gene expression (**Supplementary Table 9**). Since the aim of SMR is to prioritize variants and genes for subsequent studies, a test for heterogeneity excludes regions that may harbor multiple causal loci (pHET < 0.05).

Gene-wise and pathway analysis. Guided by rigorous method comparisons conducted by PGC members [50,65], p-values quantifying the degree of association of genes and gene sets with BD were generated using MAGMA (v1.06) [50]. We used ENSEMBL gene coordinates for 18,172 genes giving a Bonferroni corrected *P*-value threshold of 2.8×10^{-6}. Joint multi-SNP LD-adjusted gene-level p-values were calculated using SNPs 35 kb upstream to 10 kb downstream, adjusting for LD using 1,000 Genomes Project (Phase 3 v5a, MAF ≥ 0.01, European-ancestry subjects) [66]. Gene sets were compiled from multiple sources. Competitive gene set tests were conducted correcting for gene size, variant density, and LD within and between genes. The pathway map (**Supplementary Figure 6**) was constructed using the kernel generative topographic mapping algorithm (k-GTM) as described by [67]. See **Supplementary Note** for further details.

Genome build. All genomic coordinates are given in NCBI Build 37/UCSC hg19.

Availability of results. The PGC's policy is to make genome-wide summary results public. Summary statistics for our meta-analysis of the GWAS cohort samples are available through the PGC (URLs).

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

URLs

Psychiatric Genomics Consortium, PGC, https://med.unc.edu/pgc

PGC "ricopili" GWA pipeline, https://github.com/Nealelab/ricopili

1000 Genomes Project multi-ancestry imputation panel,

https://mathgen.stats.ox.ac.uk/impute/data_download_1000G_phase1_integrated.html

LD-Hub, http://ldsc.broadinstitute.org

GTEx, http://www.gtexportal.org/home/datasets

CommonMind Consortium, http://commonmind.org

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Affiliations:

1 Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, US
2 Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, US
3 Medical and Population Genetics, Broad Institute, Cambridge, MA, US
4 MRC Social, Genetic and Developmental Psychiatry Centre, King's College London, London, GB
5 NIHR BRC for Mental Health, King's College London, London, GB
6 Department of Biomedicine, University of Basel, Basel, CH
7 Department of Psychiatry (UPK), University of Basel, Basel, CH
8 Institute of Human Genetics, University of Bonn, Bonn, DE
9 Life&Brain Center, Department of Genomics, University of Bonn, Bonn, DE
10 Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, CH
11 Division of Psychiatry, University College London, London, GB
12 Stanley Center for Psychiatric Research, Broad Institute, Cambridge, MA, US
13 Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin, Berlin, DE
14 Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, US
15 iSEQ, Center for Integrative Sequencing, Aarhus University, Aarhus, DK
16 Department of Biomedicine - Human Genetics, Aarhus University, Aarhus, DK
17 Department of Clinical Neuroscience, Centre for Psychiatry Research, Karolinska Institutet, Stockholm, SE
18 Department of Psychiatry, Psychosomatics and Psychotherapy, Center of Mental Health, University Hospital Würzburg, Würzburg, DE
19 iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, DK
20 Institute of Biological Psychiatry, Mental Health Centre Sct. Hans, Copenhagen, DK
21 Institute of Clinical Medicine, University of Oslo, Oslo, NO
22 Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam, NL
23 deCODE Genetics / Amgen, Reykjavik, IS
24 Queensland Brain Institute, The University of Queensland, Brisbane, QLD, AU
25 Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD, AU
26 Division of Endocrinology and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, MA, US
27 Medical Research Council Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, GB
28 National Centre for Register-Based Research, Aarhus University, Aarhus, DK
29 Centre for Integrated Register-based Research, Aarhus University, Aarhus, DK
30 Molecular & Behavioral Neuroscience Institute, University of Michigan, Ann Arbor, MI, US
31 NEUROSCIENCE, Istituto Di Ricerche Farmacologiche Mario Negri, Milano, IT
32 Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL, US
33 Psychiatry, Berkshire Healthcare NHS Foundation Trust, Bracknell, GB
34 Psychiatry, Rush University Medical Center, Chicago, IL, US
35 Center for Neonatal Screening, Department for Congenital Disorders, Statens Serum Institut, Copenhagen, DK
36 Department of Psychiatry, Weill Cornell Medical College, New York, NY, US
37 Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, DE
38 Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, SE
39 Department of Psychiatric Research, Diakonhjemmet Hospital, Oslo, NO
40 Psychiatry, UMC Utrecht Hersencentrum Rudolf Magnus, Utrecht, NL
41 Human Genetics, University of California Los Angeles, Los Angeles, CA, US
42 Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Munich, DE

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

43 Department of Psychiatry and Human Behavior, University of California, Irvine, Irvine, CA, US
44 Molecular & Behavioral Neuroscience Institute and Department of Computational Medicine & Bioinformatics, University of Michigan, Ann Arbor, MI, US
45 Psychiatry, University of California San Francisco, San Francisco, CA, US
46 Instituto de Salud Carlos III, Biomedical Network Research Centre on Mental Health (CIBERSAM), Madrid, ES
47 Department of Psychiatry, Hospital Universitari Vall d´Hebron, Barcelona, ES
48 Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona, Barcelona, ES
49 Psychiatric Genetics Unit, Group of Psychiatry Mental Health and Addictions, Vall d´Hebron Research Institut (VHIR), Universitat Autònoma de Barcelona, Barcelona, ES
50 Department of Psychiatry, Mood Disorders Program, McGill University Health Center, Montreal, QC, CA
51 Division of Psychiatry, University of Edinburgh, Edinburgh, GB
52 University of Iowa Hospitals and Clinics, Iowa City, IA, US
53 Translational Genomics, USC, Phoenix, AZ, US
54 Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, DE
55 Centre for Psychiatry, Queen Mary University of London, London, GB
56 UCL Genetics Institute, University College London, London, GB
57 Department of Psychiatry, Laboratory of Psychiatric Genetics, Poznan University of Medical Sciences, Poznan, PL
58 Department of Neurosciences, University of California San Diego, La Jolla, CA, US
59 Department of Radiology, University of California San Diego, La Jolla, CA, US
60 Department of Psychiatry, University of California San Diego, La Jolla, CA, US
61 Department of Cognitive Science, University of California San Diego, La Jolla, CA, US
62 Applied Molecular Genomics Unit, VIB Department of Molecular Genetics, University of Antwerp, Antwerp, Belgium
63 Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, US
64 Department of Medical Genetics, Oslo University Hospital Ullevål, Oslo, NO
65 NORMENT, KG Jebsen Centre for Psychosis Research, Department of Clinical Science, University of Bergen, Bergen, NO
66 Department of Neurology, Oslo University Hospital, Oslo, NO
67 NORMENT, KG Jebsen Centre for Psychosis Research, Oslo University Hospital, Oslo, NO
68 Center for Statistical Genetics and Department of Biostatistics, University of Michigan, Ann Arbor, MI, US
69 Department of Medical & Molecular Genetics, Indiana University, Indianapolis, IN, US
70 Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, DE
71 Center for Neurobehavioral Genetics, University of California Los Angeles, Los Angeles, CA, US
72 Department of Molecular Medicine and Surgery, Karolinska Institutet and Center for Molecular Medicine, Karolinska University Hospital, Stockholm, SE
73 Department of Clinical Neuroscience, Karolinska Institutet and Center for Molecular Medicine, Karolinska University Hospital, Stockholm, SE
74 Child and Adolescent Psychiatry Research Center, Stockholm, SE
75 Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen, DE
76 Department of Psychiatry, Dalhousie University, Halifax, NS, CA
77 Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, QLD, AU
78 Department of Psychological Medicine, University of Worcester, Worcester, GB
79 School of Biomedical and Healthcare Sciences, Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, GB
80 School of Psychiatry, University of New South Wales, Sydney, NSW, AU
81 Bioinformatics Research Centre, Aarhus University, Aarhus, DK
82 Biostatistics, University of Minnesota System, Minneapolis, MN, US
83 Mental Health Department, University Regional Hospital, Biomedicine Institute (IBIMA), Málaga, ES

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

84 Department of Psychology, Eberhard Karls Universität Tübingen, Tubingen, DE
85 Department of Psychiatry and Behavioral Sciences, Howard University Hospital, Washington, DC, US
86 Center for Multimodal Imaging and Genetics, University of California San Diego, La Jolla, CA, US
87 Psychiatrie Translationnelle, Inserm U955, Créteil, FR
88 Faculté de Médecine, Université Paris Est, Créteil, FR
89 Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, CA
90 Neurogenetics Section, Centre for Addiction and Mental Health, Toronto, ON, CA
91 Department of Psychiatry, University of Toronto, Toronto, ON, CA
92 Institute of Medical Sciences, University of Toronto, Toronto, ON, CA
93 Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt, Frankfurt am Main, DE
94 Cell Biology, SUNY Downstate Medical Center College of Medicine, Brooklyn, NY, US
95 Institute for Genomic Health, SUNY Downstate Medical Center College of Medicine, Brooklyn, NY, US
96 Center for Research in Environmental Epidemiology (CREAL), Barcelona, ES
97 Psychiatry, Altrecht, Utrecht, NL
98 Psychiatry, GGZ inGeest, Amsterdam, NL
99 Psychiatry, VU medisch centrum, Amsterdam, NL
100 Psychiatry, North East London NHS Foundation Trust, Ilford, GB
101 Clinic for Psychiatry and Psychotherapy, University Hospital Cologne, Cologne, DE
102 Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA, US
103 HudsonAlpha Institute for Biotechnology, Huntsville, AL, US
104 Department of Human Genetics, University of Michigan, Ann Arbor, MI, US
105 Psychiatry, University of Illinois at Chicago College of Medicine, Chicago, IL, US
106 Max Planck Institute of Psychiatry, Munich, DE
107 Mental Health, NHS 24, Glasgow, GB
108 Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, GB
109 Psychiatry, Brigham and Women's Hospital, Boston, MA, US
110 Department of Psychiatry and Psychotherapy, University of Bonn, Bonn, DE
111 Department of Genetics, Harvard Medical School, Boston, MA, US
112 Department of Psychiatry, University of Michigan, Ann Arbor, MI, US
113 Genetic Cancer Susceptibility Group, International Agency for Research on Cancer, Lyon, FR
114 Estonian Genome Center, University of Tartu, Tartu, EE
115 Discipline of Biochemistry, Neuroimaging and Cognitive Genomics (NICOG) Centre, National University of Ireland, Galway, Galway, IE
116 Neuropsychiatric Genetics Research Group, Dept of Psychiatry and Trinity Translational Medicine Institute, Trinity College Dublin, Dublin, IE
117 Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, DE
118 Research/Psychiatry, Veterans Affairs San Diego Healthcare System, San Diego, CA, US
119 Department of Clinical Sciences, Psychiatry, Umeå University Medical Faculty, Umeå, SE
120 Department of Clinical Psychiatry, Psychiatry Clinic, Clinical Center University of Sarajevo, Sarajevo, BA
121 Department of Neurobiology, Care sciences, and Society, Karolinska Institutet and Center for Molecular Medicine, Karolinska University Hospital, Stockholm, SE
122 Psychiatry, Harvard Medical School, Boston, MA, US
123 Division of Clinical Research, Massachusetts General Hospital, Boston, MA, US
124 Outpatient Clinic for Bipolar Disorder, Altrecht, Utrecht, NL
125 Department of Psychiatry, Washington University in Saint Louis, Saint Louis, MO, US
126 Department of Biochemistry and Molecular Biology II, Institute of Neurosciences, Center for Biomedical Research, University of Granada, Granada, ES
127 Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, US
128 Medicine, Psychiatry, Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, US
129 Department of Health Sciences Research, Mayo Clinic, Rochester, MN, US

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

130 Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, US
131 Rush University Medical Center, Chicago, IL, US
132 Scripps Translational Science Institute, La Jolla, CA, US
133 Neuroscience Research Australia, Sydney, NSW, AU
134 Faculty of Medicine, Department of Psychiatry, School of Health Sciences, University of Iceland, Reykjavik, IS
135 Div Mental Health and Addiction, Oslo University Hospital, Oslo, NO
136 NORMENT, University of Oslo, Oslo, NO
137 Psychiatry and the Behavioral Sciences, University of Southern California, Los Angeles, CA, US
138 Mood Disorders, PsyQ, Rotterdam, NL
139 Institute for Medical Sciences, University of Aberdeen, Aberdeen, UK
140 Research Division, Federal Institute for Drugs and Medical Devices (BfArM), Bonn, DE
141 Centre for Addiction and Mental Health, Toronto, ON, CA
142 Neurogenomics, TGen, Los Angeles, AZ, US
143 Psychiatry, Psychiatrisches Zentrum Nordbaden, Wiesloch, DE
144 Computational Sciences Center of Emphasis, Pfizer Global Research and Development, Cambridge, MA, US
145 Department of Biostatistics, Princess Margaret Cancer Centre, Toronto, ON, CA
146 Dalla Lana School of Public Health, University of Toronto, Toronto, ON, CA
147 Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, GB
148 Department of Mental Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, US
149 Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, US
150 NORMENT, KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Institute of Clinical Medicine and Diakonhjemmet Hospital, University of Oslo, Oslo, NO
151 National Institute of Mental Health, Klecany, CZ
152 Discipline of Psychiatry, University of Adelaide, Adelaide, SA, AU
153 Department of Psychiatry and Addiction Medicine, Assistance Publique - Hôpitaux de Paris, Paris, FR
154 Paris Bipolar and TRD Expert Centres, FondaMental Foundation, Paris, FR
155 UMR-S1144 Team 1: Biomarkers of relapse and therapeutic response in addiction and mood disorders, INSERM, Paris, FR
156 Psychiatry, Université Paris Diderot, Paris, FR
157 Psychiatry, University of Pennsylvania, Philadelphia, PA, US
158 Department of Psychiatry, University of Münster, Münster, DE
159 Division of Endocrinology, Children's Hospital Boston, Boston, MA, US
160 Centre for Affective Disorders, Institute of Psychiatry, Psychology and Neuroscience, London, GB
161 Department of Psychiatry & Psychology, Mayo Clinic, Rochester, MN, US
162 School of Medical Sciences, University of New South Wales, Sydney, NSW, AU
163 Department of Human Genetics, University of Chicago, Chicago, IL, US
164 Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, RO
165 Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, SE
166 INSERM, Paris, FR
167 Department of Medical & Molecular Genetics, King's College London, London, GB
168 Neuroscience Therapeutic Area, Janssen Research and Development, LLC, Titusville, NJ, US
169 Cancer Epidemiology and Prevention, M. Sklodowska-Curie Cancer Center and Institute of Oncology, Warsaw, PL
170 School of Psychology, The University of Queensland, Brisbane, QLD, AU
171 Research Institute, Lindner Center of HOPE, Mason, OH, US
172 Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, GB
173 Human Genetics Branch, Intramural Research Program, National Institute of Mental Health, Bethesda, MD, US

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

174 Division of Mental Health and Addiction, Oslo University Hospital, Oslo, NO
175 Division of Mental Health and Addiction, University of Oslo, Institute of Clinical Medicine, Oslo, NO
176 Institute of Molecular and Cell Biology, University of Tartu, Tartu, EE
177 Mental Health, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology - NTNU, Trondheim, NO
178 Psychiatry, St Olavs University Hospital, Trondheim, NO
179 Psychosis Research Unit, Aarhus University Hospital, Risskov, DK
180 Munich Cluster for Systems Neurology (SyNergy), Munich, DE
181 University of Liverpool, Liverpool, GB
182 Psychiatry and Human Genetics, University of Pittsburgh, Pittsburgh, PA, US
183 Mental Health Services in the Capital Region of Denmark, Mental Health Center Copenhagen, University of Copenhagen, Copenhagen, DK
184 Division of Psychiatry, Haukeland Universitetssjukehus, Bergen, NO
185 Faculty of Medicine and Dentistry, University of Bergen, Bergen, NO
186 Human Genetics and Computational Biomedicine, Pfizer Global Research and Development, Groton, CT, US
187 College of Medicine Institute for Genomic Health, SUNY Downstate Medical Center College of Medicine, Brooklyn, NY, US
188 Department of Clinical Genetics, Amsterdam Neuroscience, Vrije Universiteit Medical Center, Amsterdam, NL
189 Department of Neurology and Neurosurgery, McGill University, Faculty of Medicine, Montreal, QC, CA
190 Montreal Neurological Institute and Hospital, Montreal, QC, CA
191 Department of Biomedical and NeuroMotor Sciences, University of Bologna, Bologna, IT
192 Department of Psychiatry, Massachusetts General Hospital, Boston, MA, US
193 Psychiatric and Neurodevelopmental Genetics Unit (PNGU), Massachusetts General Hospital, Boston, MA, US
194 Faculty of Medicine, University of Iceland, Reykjavik, IS
195 Department of Psychiatry, Hospital Namsos, Namsos, NO
196 Department of Neuroscience, Norges Teknisk Naturvitenskapelige Universitet Fakultet for naturvitenskap og teknologi, Trondheim, NO
197 Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, US
198 Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, US
199 Department of Psychiatry, McGill University, Montreal, QC, CA
200 Dept of Psychiatry, Sankt Olavs Hospital Universitetssykehuset i Trondheim, Trondheim, NO
201 Clinical Institute of Neuroscience, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, ES
202 Institute of Biological Psychiatry, MHC Sct. Hans, Mental Health Services Copenhagen, Roskilde, DK
203 Department of Clinical Medicine, University of Copenhagen, Copenhagen, DK
204 Psychiatry, Indiana University School of Medicine, Indianapolis, IN, US
205 Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, US

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

1 **Funding:**

Study	Lead investigator	Country, Funder, Award number
PGC	P Sullivan	USA, NIMH MH109528
PGC	D Posthuma	Netherlands, Scientific Organization Netherlands, 480-05-003
PGC	D Posthuma	Dutch Brain Foundation and the VU University Amsterdam Netherlands
Analysis, UK - BDRN (Cardiff)	PA Holmans	Medical Research Council (MRC) Centre (G0801418) and Program Grants (G0800509)
Analysis	NR Wray	NHMRC 1078901,108788
BACCS	G Breen	GB, JRIC, HG, CL were supported in part by the NIHR Maudsley Biomedical Research Centre ('BRC') hosted at King's College London and South London and Maudsley NHS Foundation Trust, and funded by the National Institute for Health Research under its Biomedical Research Centres funding initiative.
BD_TRS	U Dannlowski	Germany, DFG, Grant FOR2107 DA1151/5-1; Grant SFB-TRR58, Project C09
BiGS, Uchicago	ES Gershon	R01 MH103368
BiGS, NIMH	FJ McMahon	US, NIMH, R01 MH061613, ZIA MH002843
BiGS, GAIN, UCSD	J Kelsoe	US, NIMH, MH078151, MH081804, MH59567
BiGS, University of Pittsburgh	V Nimgaonkar	US, NIMH MH63480
BOMA-Australia	JM Fullerton	Australia, National Health and Medical Research Council, grant numbers: 1037196; 1066177; 1063960
BOMA-Australia	SE Medland	Australia, National Health and Medical Research Council, grant numbers: 1103623

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

BOMA-Australia	PB Mitchell	Australia, National Health and Medical Research Council, grant numbers: 1037196
BOMA-Australia	GW Montgomery	Australia, National Health and Medical Research Council, grant numbers: 1078399
BOMA-Australia	PR Schofield	Australia, National Health and Medical Research Council, grant numbers: 1037196
BOMA-Romania	M Grigoroiu-Serbanescu	Romania, UEFISCDI, Grant no. 89/2012
BOMA-Germany I, II, III	S Cichon	Germany, BMBF Integrament, 01ZX1314A/01ZX1614A
BOMA-Germany I, II, III	S Cichon	Germany, BMBF NGFNplus MooDS, 01GS08144
BOMA-Germany I, II, III	S Cichon	Switzerland, SNSF, 156791
BOMA-Germany I, II, III	MM Nöthen	Germany, BMBF Integrament, 01ZX1314A/01ZX1614A
BOMA-Germany I, II, III	MM Nöthen	Germany, BMBF NGFNplus MooDS, 01GS08144
BOMA-Germany I, II, III	MM Nöthen	Germany, Deutsche Forschungsgemeinschaft, Excellence Cluster ImmunoSensation
BOMA-Germany I, II, III	MM Nöthen	Germany, Deutsche Forschungsgemeinschaft, NO246/10-1
BOMA-Germany I, II, III	SH Witt	Germany, Deutsche Forschungsgemeinschaft, WI 3429/3-1
BOMA-Germany I, II, III, BOMA-Spain	M Rietschel	Germany, BMBF Integrament, 01ZX1314G/01ZX1614G
BOMA-Germany I, II, III, BOMA-Spain	M Rietschel	Germany, BMBF NGFNplus MooDS, 01GS08147
BOMA-Germany I, II, III, BOMA-Spain	M Rietschel	Germany, Deutsche Forschungsgemeinschaft, RI 908/11-1

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

BOMA-Germany I, II, III, PsyCourse, BiGS	TG Schulze	Germany, BMBF Integrament, 01ZX1314K
BOMA-Germany I, II, III, PsyCourse, BiGS	TG Schulze	Germany, DFG, SCHU 1603/4-1, SCHU 1603/5-1, SCHU 1603/7-1
BOMA-Germany I, II, III, PsyCourse, BiGS	TG Schulze	Germany, Dr. Lisa-Oehler Foundation (Kassel, Germany)
Bulgarian Trios (Cardiff)	G Kirov	The recruitment was funded by the Janssen Research Foundation. Genotyping was funded by multiple grants to the Stanley Center for Psychiatric Research at the Broad Institute from the Stanley Medical Research Institute, The Merck Genome Research Foundation, and the Herman Foundation.
Fran	M Leboyer	France, Inserm, ANR
Halifax	M Alda	CIHR grant #64410
iPSYCH BP group	AD Børglum	Denmark, Lundbeck Foundation, R102-A9118 and R155-2014-1724 (iPSYCH)
iPSYCH BP group	AD Børglum	Denmark, Aarhus University, iSEQ and CIRRAU
iPSYCH BP group	AD Børglum	USA, Stanley Medical Research Institute
iPSYCH BP group	AD Børglum	EU, European Research Council, 294838
Mayo Bipolar Disorder Biobank	JM Biernacka, MA Frye	Marriot Foundation and the Mayo Clinic Center for Individualized Medicine
Michigan	M Boehnke	US, NIMH, R01 MH09414501A1; US, NIMH, MH105653
Mount Sinai	EA Stahl	NARSAD Young Investigator Award
Mount Sinai, STEP-BD, FAST	P Sklar, EA Stahl	US NIH R01MH106531, R01MH109536
NeuRA-CASSI-Australia	C Shannon Weickert	Australia, National Health and Medical Research Council, grant number: 568807

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

NeuRA-CASSI-Australia	TW Weickert	Australia, National Health and Medical Research Council, grant number: 568807
NeuRA-IGP-Australia	MJ Green	Australia, National Health and Medical Research Council, grant numbers: 630471, 1081603
Norway	I Agartz	Sweden, Swedish Research Council
Norway	OA Andreassen	Norway, Research Council of Norway (#217776, #223273, #248778, #249711), KG Jebsen Stiftelsen, The South-East Norway Regional Health Authority (#2012-132, #2012-131, #2017-004)
Norway	T Elvsåshagen	Norway, The South-East Norway Regional Health Authority (#2015-078) and a research grant from Mrs. Throne-Holst.
Norway	I Melle	Norway, Research Council of Norway (#421716,#223273), KG Jebsen Stiftelsen, The South-East Norway Regional Health Authority (#2011085, #2013088, #2014102)
Norway	KJ Oedegaard	Norway, the Western Norway Regional Health Authority
Norway	OB Smeland	Norway, The South-East Norway Regional Health Authority (#2016-064, #2017-004)
Span2	M Ribasés	Spain, Instituto de Salud Carlos III, Ministerio de Economía, Industria y Competitividad, CP09/00119 and CPII15/00023
Span2	C Sánchez-Mora	Spain, Instituto de Salud Carlos III, Ministerio de Economía, Industria y Competitividad, CD15/00199 and MV16/00039
State University of New York, Downstate Medical Center (SUNY DMC)	C Pato, MT Pato, JA Knowles, H Medeiros	US, National Institutes of Health, R01MH085542
SWEBIC	M Landén	The Stanley Center for Psychiatric Research, Broad Institute from a grant from Stanley Medical Research Institute; NIMH MH077139 (PFS), The Swedish Research Council (K2014-62X-14647-12-51 and K2010-61P-21568-01-4), and the Swedish foundation for Strategic Research (KF10-0039)

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

UCL	A McQuillin	Medical Research Council (MRC) - G1000708
UCLA-Utrecht (Los Angeles)	NB Freimer	US, National Institutes of Health, R01MH090553, U01MH105578
UCLA-Utrecht (Los Angeles)	LM Olde Loohuis	US, National Institutes of Health, R01MH090553, U01MH105578
UCLA-Utrecht (Los Angeles)	RA Ophoff	US, National Institutes of Health, R01MH090553, U01MH105578
UCLA-Utrecht (Los Angeles)	APS Ori	US, National Institutes of Health, R01MH090553, U01MH105578
UK - BDRN (Cardiff)	MC O'Donovan	Medical Research Council (MRC) Centre (G0801418) and Program Grants (G0800509)
UK - BDRN (Cardiff)	MJ Owen	Medical Research Council (MRC) Centre (G0801418) and Program Grants (G0800509)
UK - BDRN (Cardiff)	N Craddock, I Jones, LA Jones	UK, Wellcome Trust, 078901; USA, Stanley Medical Research Institute, 5710002223-01
UK - BDRN (Cardiff)	A Di Florio	European Commission Marie Curie Fellowship, grant number 623932.
UNIBO / University of Barcelona, Hospital Clinic, IDIBAPS, CIBERSAM	E Vieta	Grants PI15/00283 (Spain) and 2014 SGR 398 (Catalonia)
USC	JL Sobell	USA, National Institutes of Health, R01MH085542
WTCCC	N Craddock; AH Young	The principal funder of this project was the Wellcome Trust. For the 1958 Birth Cohort, venous blood collection was funded by the UK Medical Research Council. AHY was funded by NIMH (USA); CIHR (Canada); NARSAD (USA); Stanley Medical Research Institute (USA); MRC (UK); Wellcome Trust (UK); Royal College of Physicians (Edin); BMA (UK); UBC-VGH Foundation (Canada); WEDC (Canada); CCS Depression Research Fund (Canada); MSFHR (Canada); NIHR (UK); Janssen (UK)

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Acknowledgments:

This paper is dedicated to the memory of Psychiatric Genomics Consortium (PGC) founding member and Bipolar disorder working group co-chair Pamela Sklar. We are deeply indebted to the investigators who comprise the PGC, and to the subjects who have shared their life experiences with PGC investigators. The PGC has received major funding from the US National Institute of Mental Health (PGC3: U01 MH109528, PGC2: U01 MH094421, PGC1: U01 MH085520). Statistical analyses were carried out on the NL Genetic Cluster Computer (http://www.geneticcluster.org) hosted by SURFsara.

BACCS: This work was supported in part by the NIHR Maudsley Biomedical Research Centre ('BRC') hosted at King's College London and South London and Maudsley NHS Foundation Trust, and funded by the National Institute for Health Research under its Biomedical Research Centres funding initiative. The views expressed are those of the authors and not necessarily those of the BRC, the NHS, the NIHR or the Department of Health or King's College London. We gratefully acknowledge capital equipment funding from the Maudsley Charity (Grant Reference 980) and Guy's and St Thomas's Charity (Grant Reference STR130505).

BD_TRS: This work was funded by the German Research Foundation (DFG, grant FOR2107 DA1151/5-1 to UD; SFB-TRR58, Project C09 to UD) and the Interdisciplinary Center for Clinical Research (IZKF) of the medical faculty of Münster (grant Dan3/012/17 to UD).

BiGS, GAIN: FJM was supported by the NIMH Intramural Research Program, NIH, DHHS.

BOMA-Australia: JMF would like to thank Janette M O'Neil and Betty C Lynch for their support.

BOMA-Germany I, BOMA-Germany II, BOMA-Germany III, PsyCourse: This work was supported by the German Ministry for Education and Research (BMBF) through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med program (grant 01ZX1314A/01ZX1614A to MMN and SC, grant

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

01ZX1314G/01ZX1614G to MR, grant 01ZX1314K to TGS). This work was supported by the German Ministry for Education and Research (BMBF) grants NGFNplus MooDS (Systematic Investigation of the Molecular Causes of Major Mood Disorders and Schizophrenia; grant 01GS08144 to MMN and SC, grant 01GS08147 to MR). This work was also supported by the Deutsche Forschungsgemeinschaft (DFG), grant NO246/10-1 to MMN (FOR 2107), grant RI 908/11-1 to MR (FOR 2107), grant WI 3429/3-1 to SHW, grants SCHU 1603/4-1, SCHU 1603/5-1 (KFO 241) and SCHU 1603/7-1 (PsyCourse) to TGS. This work was supported by the Swiss National Science Foundation (SNSF, grant 156791 to SC). MMN is supported through the Excellence Cluster ImmunoSensation. TGS is supported by an unrestricted grant from the Dr. Lisa-Oehler Foundation. AJF received support from the BONFOR Programme of the University of Bonn, Germany. MH was supported by the Deutsche Forschungsgemeinschaft.

Edinburgh: DJM is supported by an NRS Clinical Fellowship funded by the CSO.

Fran: This research was supported by Foundation FondaMental, Créteil, France and by the Investissements d'Avenir Programs managed by the ANR under references ANR-11-IDEX-0004-02 and ANR-10-COHO-10-01.

Halifax: Halifax data were obtained with support from the Canadian Institutes of Health Research.

iPSYCH BP group: ADB and the iPSYCH team acknowledges funding from The Lundbeck Foundation (grant no R102-A9118 and R155-2014-1724), the Stanley Medical Research Institute, an Advanced Grant from the European Research Council (project no: 294838), and grants from Aarhus University to the iSEQ and CIRRAU centers.

The Mayo Bipolar Disorder Biobank was funded by the Marriot Foundation and the Mayo Clinic Center for Individualized Medicine

Michigan (NIMH/Pritzker Neuropsychiatric Disorders Research Consortium): We thank the

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

participants who donated their time and DNA to make this study possible. We thank members of the NIMH Human Genetics Initiative and the University of Michigan Prechter Bipolar DNA Repository for generously providing phenotype data and DNA samples. Many of the authors are members of the Pritzker Neuropsychiatric Disorders Research Consortium which is supported by the Pritzker Neuropsychiatric Disorders Research Fund L.L.C. A shared intellectual property agreement exists between this philanthropic fund and the University of Michigan, Stanford University, the Weill Medical College of Cornell University, HudsonAlpha Institute of Biotechnology, the Universities of California at Davis, and at Irvine, to encourage the development of appropriate findings for research and clinical applications.

Mount Sinai: This work was funded in part by a NARSAD Young Investigator award to EAS.

NeuRA-CASSI-Australia: This work was funded by the NSW Ministry of Health, Office of Health and Medical Research. CSW was a recipient of National Health and Medical Research Council (Australia) Fellowships (#1117079, #1021970).

NeuRA-IGP-Australia: MJG was supported by a NHMRC Career Development Fellowship (1061875).

Norway: TE was funded by The South-East Norway Regional Health Authority (#2015-078) and a research grant from Mrs. Throne-Holst.

Span2: CSM is a recipient of a Sara Borrell contract (CD15/00199) and a mobility grant (MV16/00039) from the Instituto de Salud Carlos III, Ministerio de Economía, Industria y Competitividad, Spain. MR is a recipient of a Miguel de Servet contract (CP09/00119 and CPII15/00023) from the Instituto de Salud Carlos III, Ministerio de Economía, Industria y Competitividad, Spain. This investigation was supported by Instituto de Salud Carlos III (PI12/01139, PI14/01700, PI15/01789, PI16/01505), and cofinanced by the European Regional Development Fund (ERDF), Agència de Gestió d'Ajuts Universitaris i de Recerca-AGAUR,

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Generalitat de Catalunya (2014SGR1357), Departament de Salut, Generalitat de Catalunya, Spain, and a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation. This project has also received funding from the European Union's Horizon 2020 Research and Innovation Programme under the grant agreements No 667302 and 643051.

SWEBIC: We are deeply grateful for the participation of all subjects contributing to this research, and to the collection team that worked to recruit them. We also wish to thank the Swedish National Quality Register for Bipolar Disorders: BipoläR. Funding support was provided by the Stanley Center for Psychiatric Research, Broad Institute from a grant from Stanley Medical Research Institute, the Swedish Research Council, and the NIMH.

Sweden: This work was funded by the Swedish Research Council (M. Schalling, C. Lavebratt), the Stockholm County Council (M. Schalling, C. Lavebratt, L. Backlund, L. Frisén, U. Ösby) and the Söderström Foundation (L. Backlund).

UK - BDRN: BDRN would like to acknowledge funding from the Wellcome Trust and Stanley Medical Research Institute, and especially the research participants who continue to give their time to participate in our research.

UNIBO / University of Barcelona, Hospital Clinic, IDIBAPS, CIBERSAM: EV thanks the support of the Spanish Ministry of Economy and Competitiveness (PI15/00283) integrated into the Plan Nacional de I+D+I y cofinanciado por el ISCIII-Subdirección General de Evaluación y el Fondo Europeo de Desarrollo Regional (FEDER); CIBERSAM; and the Comissionat per a Universitats i Recerca del DIUE de la Generalitat de Catalunya to the Bipolar Disorders Group (2014 SGR 398).

WTCCC: The principal funder of this project was the Wellcome Trust. For the 1958 Birth Cohort, venous blood collection was funded by the UK Medical Research Council.

AHY is funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Conflicts of Interest:

T.E. Thorgeirsson, S. Steinberg, H. Stefansson and K. Stefansson are employed by deCODE Genetics/Amgen. Multiple additional authors work for pharmaceutical or biotechnology companies in a manner directly analogous to academic co-authors and collaborators. A.H. Young has given paid lectures and is on advisory boards for the following companies with drugs used in affective and related disorders: Astrazenaca, Eli Lilly, Janssen, Lundbeck, Sunovion, Servier, Livanova. A.H. Young is Lead Investigator for Embolden Study (Astrazenaca), BCI Neuroplasticity study and Aripiprazole Mania Study, which are investigator-initiated studies from Astrazenaca, Eli Lilly, Lundbeck, and Wyeth. J. Nurnberger is an investigator for Janssen. P.F. Sullivan reports the following potentially competing financial interests: Lundbeck (advisory committee), Pfizer (Scientific Advisory Board member), and Roche (grant recipient, speaker reimbursement). G. Breen reports consultancy and speaker fees from Eli Lilly and Illumina and grant funding from Eli Lilly. O.A. Andreassen has received speaker fees from Lundbeck. All other authors declare no financial interests or potential conflicts of interest.

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

DISPLAY ITEMS (inline above in this manuscript version):

Figure 1. Manhattan plot for our primary genomewide association analysis of 20,352 cases and 31,358 controls. GWAS -$\log_{10}$P-values are plotted for all SNPs across chromosomes 1-22 (diamonds, green for loci with lead SNP GWAS P < 10^{-6}). Combined GWAS+followup -$\log_{10}$P-values for lead SNPs reaching genome-wide significance in either GWAS or combined analysis (triangles, inverted if GWAS+followup -$\log_{10}$P > GWAS -$\log_{10}$P). Labels correspond to gene symbols previously reported for published loci (black) and the nearest genes for novel loci (blue), at top if GWAS+followup P < $5x10^{-8}$.

Table 1. Genome-wide significant bipolar disorder risk loci.

Figure 2. Association of BD1 and BD2 subtypes with schizophrenia (SCZ) and major depression (DEPR) polygenic risk scores (PRS). Shown are mean PRS values (1 s.e. error bars), adjusted for study and ancestry covariates and scaled to the PRS mean and sd in control subjects, in BD1 (red) and BD2 (blue) cases, for increasing source GWAS P-value thresholds (increasing grey) as indicated. P-values (italics) test BD1 vs BD2 mean PRS, in logistic regression of case subtype on PRS with covariates. Results are detailed in Supplementary Table 10.

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

References:

1. Ferrari, A. J. *et al.* The prevalence and burden of bipolar disorder: findings from the Global Burden of Disease Study 2013. *Bipolar Disord.* **18,** 440–450 (2016).
2. Lichtenstein, P. *et al.* Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* **373,** 234–239 (2009).
3. Edvardsen, J. *et al.* Heritability of bipolar spectrum disorders. Unity or heterogeneity? *J. Affect. Disord.* **106,** 229–240 (2008).
4. McGuffin, P. *et al.* The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch. Gen. Psychiatry* **60,** 497–502 (2003).
5. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. (American Psychiatric Pub, 2013).
6. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. (World Health Organization, 1992).
7. Craddock, N. & Owen, M. J. The Kraepelinian dichotomy - going, going... but still not gone. *Br. J. Psychiatry* **196,** 92–95 (2010).
8. Lee, S. H. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45,** 984–994 (2013).
9. Sklar, P. *et al.* Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat. Genet.* **43,** 977–U162 (2011).
10. 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).
11. Charney, A. W. *et al.* Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder. *Transl. Psychiatry* **7,** e993 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

12. Chen, D. T. *et al.* Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Mol. Psychiatry* **18,** 195–205 (2013).

13. Cichon, S. *et al.* Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am. J. Hum. Genet.* **88,** 372–381 (2011).

14. Ferreira, M. A. R. *et al.* Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat. Genet.* **40,** 1056–1058 (2008).

15. Green, E. K. *et al.* Association at SYNE1 in both bipolar disorder and recurrent major depression. *Mol. Psychiatry* **18,** 614–617 (2013).

16. Green, E. K. *et al.* Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. *Mol. Psychiatry* **18,** 1302–1307 (2013).

17. Hou, L. *et al.* Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. *Hum. Mol. Genet.* **25,** 3383–3394 (2016).

18. Mühleisen, T. W. *et al.* Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat. Commun.* **5,** 3339 (2014).

19. Schulze, T. G. *et al.* Two variants in Ankyrin 3 (ANK3) are independent genetic risk factors for bipolar disorder. *Mol. Psychiatry* **14,** 487–491 (2009).

20. Scott, L. J. *et al.* Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 7501–7506 (2009).

21. Sklar, P. *et al.* Whole-genome association study of bipolar disorder. *Mol. Psychiatry* **13,** 558–569 (2008).

22. Smith, E. N. *et al.* Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol. Psychiatry* **14,** 755–763 (2009).

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

23. Burton, P. R. *et al.* Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447,** 661–678 (2007).

24. Gratten, J., Wray, N. R., Keller, M. C. & Visscher, P. M. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat. Neurosci.* **17,** 782–790 (2014).

25. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).

26. Palmer, C. & Pe'er, I. Statistical correction of the Winner's Curse explains replication variability in quantitative trait genome-wide association studies. *PLoS Genet.* **13,** e1006916 (2017).

27. Zhong, H. & Prentice, R. L. Bias-reduced estimators and confidence intervals for odds ratios in genome-wide association studies. *Biostatistics* **9,** 621–634 (2008).

28. Gao, X., Starmer, J. & Martin, E. R. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.* **32,** 361–369 (2008).

29. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511,** 421–427 (2014).

30. Goes, F. S. *et al.* Genome-wide association study of schizophrenia in Ashkenazi Jews. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **168,** 649–659 (2015).

31. Ripke, S. *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat. Genet.* **45,** 1150–1159 (2013).

32. Wray, N. R. & Sullivan, P. F. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *bioRxiv* (2017).

33. Okbay, A. *et al.* Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* (2016).

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

doi:10.1038/ng.3552

34. Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45,** 984–994 (2013).

35. Duncan, L. *et al.* Significant Locus and Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa. *Am. J. Psychiatry* appiajp201716121402 (2017).

36. Otowa, T. *et al.* Meta-analysis of genome-wide association studies of anxiety disorders. *Mol. Psychiatry* **21,** 1391–1399 (2016).

37. Gale, C. R. *et al.* Pleiotropy between neuroticism and physical and mental health: findings from 108 038 men and women in UK Biobank. *Transl. Psychiatry* **6,** e791 (2016).

38. Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* **340,** 1467–1471 (2013).

39. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533,** 539–542 (2016).

40. Benyamin, B. *et al.* Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. *Mol. Psychiatry* **19,** 253–258 (2014).

41. Sniekers, S. *et al.* Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat. Genet.* **49,** 1107–1112 (2017).

42. Hou, L. *et al.* Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. *Hum. Mol. Genet.* **25,** 3383–3394 (2016).

43. Ikeda, M. *et al.* A genome-wide association study identifies two novel susceptibility loci and trans population polygenicity associated with bipolar disorder. *Mol. Psychiatry* (2017). doi:10.1038/mp.2016.259

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

1 44. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex
2 trait gene targets. *Nat. Genet.* **48,** 481–487 (2016).
3 45. Fromer, M. *et al.* Gene expression elucidates functional impact of polygenic risk for
4 schizophrenia. *Nat. Neurosci.* **19,** 1442–1453 (2016).
5 46. Westra, H. J. *et al.* Systematic identification of trans eQTLs as putative drivers of known
6 disease associations. *Nat. Genet.* **45,** 1238–1243 (2013).
7 47. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide
8 association summary statistics. *Nat. Genet.* **47,** 1228–1235 (2015).
9 48. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human
10 epigenomes. *Nature* **518,** 317–330 (2015).
11 49. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using
12 predicted gene functions. *Nat. Commun.* **6,** 5890 (2015).
13 50. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set
14 analysis of GWAS data. *PLoS Comput. Biol.* **11,** e1004219 (2015).
15 51. Palmer, C. & Pe'er, I. Statistical Correction of the Winner9s Curse Explains Replication
16 Variability in Quantitative Trait Genome-Wide Association Studies. *bioRxiv* 104786 (2017).
17 52. Gaspar, H. A. & Breen, G. Pathways analyses of schizophrenia GWAS focusing on known
18 and novel drug targets. *Biorxiv* (2017). doi:10.1101/091264
19 53. Camandola, S. & Mattson, M. P. Aberrant subcellular neuronal calcium regulation in aging
20 and Alzheimer's disease. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*
21 **1813,** 965–973 (2011).
22 54. Mertens, J. *et al.* Differential responses to lithium in hyperexcitable neurons from patients
23 with bipolar disorder. *Nature* **527,** 95–99 (2015).
24 55. Lee, S. H., Goddard, M. E., Wray, N. R. & Visscher, P. M. A better coefficient of

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

determination for genetic profile analysis. *Genet. Epidemiol.* **36,** 214–224 (2012).

56. Nolen, W. A. The continuum of unipolar depression - bipolar II depression - bipolar I depression: different treatments indicated? *World Psychiatry* **10,** 196–197 (2011).
57. MacCabe, J. H. *et al.* Excellent school performance at age 16 and risk of adult bipolar disorder: national cohort study. *Br. J. Psychiatry* **196,** 109–115 (2010).
58. Vreeker, A. *et al.* High educational performance is a distinctive feature of bipolar disorder: a study on cognition in bipolar disorder, schizophrenia patients, relatives and controls. *Psychol. Med.* **46,** 807–818 (2016).
59. Ripke, S. Ricopili: a tool for visualizing regions of interest in select GWAS data sets. (2014).
60. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
61. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4,** 7 (2015).
62. Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31,** 1466–1468 (2015).
63. Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33,** 272–279 (2017).
64. Finucane, H. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. doi:10.1101/103069
65. O'Dushlaine, C. *et al.* Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat. Neurosci.* **18,** 199–209 (2015).
66. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526,** 68–74 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/173062. this version posted January 24, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

67. Olier, I., Vellido, A. & Giraldo, J. Kernel generative topographic mapping. in *ESANN* **2010,** 481–486 (2010).