Instructions: Complete this template to provide IRB members and designated reviewers with sufficient information to conduct a substantive review of human research. If applicable, submit a Sponsor's Protocol in addition to this document. Detailed instructions for preparing this template can be found in the Investigator's Manual. If the proposed human research is eligible for an Exemption Determination, see Appendix H of the Investigator Manual.

GENERAL INFORMATION	
Protocol/ESTR Record Number (if assigned): IRB20-1510	
Version Number: 4.0	**Version Date: 2/27/2023**
Principal Investigator (PI): Hailiang Huang	
Principal Investigator's Harvard Affiliation: Faculty	
Protocol Title: Asian Bipolar Genetics Network (A-BIG-NET)	

1. Specific Aims

Aim 1. Recruitment and phenotype quality control alignment. a) We will recruit a total of 19,500 Bipolar I Disorder (BP-I) cases and 15,000 controls from five countries in Asia. A rich array of phenotyping information will be collected including demographic, environmental and detailed clinical and historical features. Cross-site consistency will be assured and aided by a uniform cross-site data collection protocol. Quality control procedures will be implemented.
Outcome: DNA samples and deep phenotyping data from 19,500 BP-I cases and 15,000 controls.
Aim 2: Genetics data production and quality control. a) We will use a novel and cost effective genomic technology combining 1xWGS and 55xWES, developed by the Broad Institute Genomics Platform, on all recruited samples and additional samples from a funded bipolar genetics study of Pakistanian populations (10,000 cases and 2,000 controls; Project Number: 1R01MH123775). **b)** DNA samples from India and South Korea cannot be shipped to the U.S. per export restrictions. We will work with their genomic facilities to use the same technology to minimize batch effects. **c)** We will perform alignment, variant calling, genotype refinement and quality control uniformly on the genetics data generated from the Broad Institute, India and South Korea.
Outcome: Genomics data, including 1xWGS and 55xWES, on 27,500 BP cases and 16,000 controls, will be shared with collaborators through Terra (Terra is a cloud-native platform for biomedical researchers to access data, run analysis tools, and collaborate) and dbGAP (with the broader scientific community).
Aim 3: Comprehensive analyses on the phenotyping and genetic data from Aims 1 and 2. **a)** We will perform studies to identify genetic variants associated with BP-I in the Asian populations and explore phenotypic variance in BP-I and its relationship with genetic signatures and environmental risk factors. **b)** We will contribute our data to the Psychiatric Genomics Consortium and Bipolar Sequencing Project, and jointly analyze with their BP-I data of European ancestry for the comparative genetic architecture of BP-I and the transferability of BP-I polygenic risk score across populations. **c)** we will jointly analyze the BP-I and SCZ samples of Asian ancestry to dissect their cross-disorder genetic architecture in the Asian populations. **d)** we will perform a cross-ancestry, cross-disease fine-mapping analysis to identify putative causal variants for further characterization of the molecular mechanisms underlying severe psychiatric disorders .
Outcome: Knowledge in the BP-I genetics within Asia and across Asia and European populations. All results will be promptly and shared through web portals and as downloadable files.
This proposed study will dramatically increase the diversity of genetic discovery efforts by creating a BP-I genetics resource with rich phenotyping data, an important step towards reducing health

disparities and accelerate gene discovery for psychiatric disorders in cohorts of diverse ancestry to advance global mental health discovery and equity.

2. Background and Significance

2.1 Provide the scientific background and rationale for the research.

Bipolar disorder (BP) is a severe multifactorial neuropsychiatric disorder with a life-time prevalence of 1-2%. A recent large-scale genetic study identified 30 BP associated genetic loci, providing initial insights into BP pathogenesis. Nevertheless, **genetic discovery in BP lags behind other key psychiatric disorders**: the reported BP genetic loci only capture a small proportion of the total BP genetic liability, with many more across common and rare allele frequency spectrum, remaining to be discovered. In addition, **all samples in the study were of European ancestry (EUR)**, leaving population specific BP variants uncovered and the uncertainty in how the BP genetic findings can be applicable to other populations, exacerbating healthcare disparities.

In response to NIMH PAR-19-297 grant (not awarded yet), we propose the Asian Bipolar Genetics Network (A-BIG-NET), an international collaboration of investigators, many with prior collaborative experience and co-publication, from the U.S., Japan, South Korea, Taiwan, Singapore, India and Pakistan. All have strong track records of large-scale psychiatric genetic research in Asia populations. Our team also contains several of the field's leaders in genetic field studies and analyses. **Studying BP genetics in Asia is important to the world and the U.S**., as Asia constitutes 57% of the world population, and many underrepresented groups in the U.S., such as Native Americans and Hispanics, descend in part from early Asian populations. The six countries in A-BIG-NET cover 40% of the Asian populations.

2.2 Describe the significance of the research, and how it will contribute to generalizable knowledge.

A-BIG-NET will generate a BP-I genetic resource of 27,500 cases and 16,000 controls with rich phenotypic information and genetics data from a novel technology combining low-pass whole genome (1xWGS) and deep exome sequencing (55xWES). We will focus on BP-I disorder to maximize genetic homogeneity. This resource is powered to discover new genetic associations with BP-I across the allelic spectrum in the Asia populations, and will reveal, for the first time, the comparative genetic architecture of BP-I across major world populations. Combining schizophrenia (SCZ) genetics resources that our team has already put together (15,000 cases and 14,000 controls), we will further expand the scope of psychiatric genetics by extending to two of the most important and severe adult psychiatric disorders in Asian populations.

This proposed study will dramatically increase the diversity of genetic discovery efforts by creating a BP-I genetics resource with rich phenotyping data, an important step towards reducing health disparities and accelerate gene discovery for psychiatric disorders in cohorts of diverse ancestry to advance global mental health discovery and equity.

3. Research Locations and Collaborating Sites

Research Locations refer to the geographic location that the research will take place, not to the institutions or researchers you may be collaborating with. All Research Locations should be listed in ESTR as a Research Location.

Collaborating Sites refer to institutions or researchers that are also taking part in the research study. All Collaborating Sites should be listed in ESTR as a Participating Site.

3.1. Where will the research activities take place? (check all that apply)

☐ **At Harvard;** *list any non-Harvard Longwood Medical Area (LMA) Schools here: A list of all Harvard Schools can be found here.*

☐ **At another location in Massachusetts;** *specify here:*

☐ **In another state in the U.S.;** *specify here:*

☒ **Internationally;** *specify here: US, Japan, South Korea, India, Singapore and Taiwan*

3.2. Describe the sites or locations where the research will be conducted or overseen by the Harvard PI. (If conducting the study virtually or remotely, indicate the location of the researcher who is conducting the study.)

Genomic data generation and analysis of genomic and phenotypic data will be conducted at the Broad Institute in Cambridge, MA. Genomic data will be generated from blood samples collected by our Asian collaborators. The study will be approved by local IRB before starting sample collection. For samples from India and Korea, genomic data generation will be conducted at facilities inside the country and will be shared with the Broad Institute. Non-identifying phenotypic data collected under this locally approved protocol will also be provided to the Broad Institute for analysis in conjunction with the genetic data.

Participant recruitment will take place in the following sites in Asia:

Japan: Japanese collections are led by Juntendo University. Other three sites joining the collection in Japan are Fujita Medical University, Kansai Medical University and Tokyo Metropolitan Institute of Medical Science.

South Korea: Korean collections are led by Korea University Hospital. A total of 12 hospitals across Korea will join for participants recruitment. The sites are: Korea University Hospitals, Seoul National University Hospital, Budang Seoul National University Hospital, Samsung Seoul Hospital, Severance Hospital, Kyungbul National University Hospital, Busan National University Hospital, Chunnam National University Hospital, National Center of Mental Health, Kyungsang University Hospital, Eulji Hosptal and Ilsan Paik Hospital.

India: Indian collections are led by National Institute of Mental Health And Neurosciences(NIMHANS). NIMHANS will be collaboration sites from IC-MAGIC consortium (https://www.med.unc.edu/pgc/about-us/people/pgc-ic-magic/) in recruitment efforts.

Singapore: Recruitment will take place in the Institute of Mental Health, Singapore.

Taiwan: National Taiwan University Hospital (NTUH) will be the coordinating center for the recruitment in Taiwan. The recruitment will take place in regional and central hospitals in the northern, central, southern and eastern of Taiwan.

3.3. Describe plans for communication among sites regarding adverse events, interim results, protocol modifications, monitoring of data, etc. ☐ N/A.

The PIs, Key Personnel and study staff from the Broad Institute, Virginia Commonwealth University, Johns Hopkins University and the National University of Taiwan will conduct monthly steering committee calls once the award period begins, and communicate via email throughout the duration of the study to share updates and information from each of the recruitment sites, phenotypic collection and harmonization efforts, sample intake, data generation, data analysis and data sharing. If site leaders are unavailable, they will ensure that a representative is on the call to participate in their place. Once we enter into the analysis phase of the grant, we will schedule bi-weekly phenotype and analysis calls, in addition to the steering committee calls.

The project manager at the Broad Institute (Zhenglin Guo), with oversight from PI Huang, will communicate on a frequent basis with the East Asia site PIs and staff members (Taiwan, Japan, South Korea and Singapore) to address in a timely manner any issues that arise in recruitment. Similarly, the lead project coordinator at JHU (Michael Morreale), with oversight from PI Zandi, will communicate in a similar fashion with the Indian sites. Issues related to phenotyping will be directed initially to the clinical psychologist trained and directly supervised by PI Kendler in the phenotyping core at VCU. If the question cannot be dealt with by that individual, then it will be triaged to one of two experienced psychiatrists: for the East Asian sites to Dr Jacob Taylor at the Broad Institute and for South Asian sites to Dr. Fernando Goes at the Johns Hopkins University. If the issues cannot be resolved at that level, it will be brought to the next regular meeting of Drs. Kendler, Taylor and Goes for a final decision. Such matters will typically be recorded in the Question by Question commentary on the interview instrument, updated versions of which will be distributed to all sites at regular intervals to ensure future instances of this problem will be dealt with in the same manner. The clinical psychologist working at VCU will also be responsible for routine checks on the completeness and quality of the phenotyping assessment and the quality controls efforts across all sites on a daily basis.

3.4. Describe any local (international or state) laws, regulations, and/or customs affecting the research (e.g., age of majority, mandatory reporting requirements, etc). ☐ N/A.

Local IRB approvals from each participating country are required prior to starting recruitment at sites. These approvals, along with their corresponding consent forms, will be submitted to the Harvard IRB for review and approval as they are issued by the local sites. Study activities in each country will not proceed until both the local IRB and the Harvard IRB have issued the necessary approvals.

For India, sharing de-identified data with the US will need approval from Indian Council of Medical Research (ICMR). We will upload the approval to the ESTR system once it's available.

In Taiwan, legal adult age is set at 20 years old. Thus, the recruitment in Taiwan will be over 20 years old.

In Singapore, legal adult age is set at 21 years old. Thus, the recruitment in Singapore will be over 21 years old.

Virginia Commonwealth University IRB decided to reliant on HSPH IRB, the reliance request is approved. They will only be involved in analyzing de-identified data. For Johns Hopkins University, as they will only working on de-identified data analysis and don't have contact with human subjects, their IRB determined that they are exempt from human subjects research. The IRB approval will be uploaded to ESTR once it's available.

Our research will be conducted follow the local regulations and laws, especially those particularly for patients with mental health conditions and for anyone who may harm themselves or others. Please see the detailed reporting requirements regarding mental health patients and study team's plan in the below table:

Country	Legal Requirements/Reporting	Mental Health Conditions, Referrals & Reporting
Japan	No legally mandatory reporting regulations for researchers. In Japan, some officers, such as policemen, are obliged to report anyone who may harm themselves or the public. Anyone who found it should try to report.	The study team will refer to the treating psychiatrist in the event of self-injury or harm to others.
India	The Mental Healthcare Act in India only requires mandatory reporting with patients who are under involuntary admissions.	The study team will refer to the treating psychiatrist in the event of self-injury or harm to others.
Korea	In Korea, there is no obligation to report when finding a mentally ill person at risk of self-injury or harm to others.	The study team will refer to the treating psychiatrist in the event of self-injury or harm to others.

Singapore	The law specific to mental health patients relates to risk of violence and self-harm/suicide. If an individual with suspected or confirmed mental illness is assessed to be at imminent risk of a violent act or suicide, researchers are obliged to notify the police. Other forms of violence, e.g. sexual violence or domestic violence, or illicit drug use, it relates to laws applicable to all individuals, regardless of mental illness, and researchers are also legally obliged to report.	The study team will notify and refer the patient to the treating psychiatrist if there's risk of violence and self-harm/suicide.
Taiwan	Two laws related to self-harm in Taiwan are 'Mental Health Act' and 'Suicide Prevention Act', and the law of 'Domestic Violence Prevention Act' for domestic violence (DV). Researchers are mandatory to report self-harm and DVs if there are suspected/confirmed cases.	The study team will notify patients' treating psychiatrist in the case of high risk of self-harm and the medical team will complete the reporting process.

3.5. Identify any approvals or permissions required of collaborating institutions, community leaders, or government officials, including approval from another IRB or local research ethics committee. *Upload copies to the "Study-Related Documents" page in ESTR.* ☐ **N/A.**

Local IRB approvals from each participating country are required prior to starting recruitment at sites. Currently, we have obtained approval from Japan Juntendo University which are uploaded in ESTR. We will upload additional local IRB approvals for the other sites as they become available.

While our entire NIMH grant is still under review, our collaborators from Singapore, Taiwan and Korea have some of their own funding and are willing to start enrollment early. Therefore, we are submitting the initial IRB application from Singapore and Taiwan for review first.

For India, sharing de-identified data with the US will need approval from Indian Council of Medical Research (ICMR). We will upload the approval to the ESTR system once it's available. Since the data will come to Broad, we will have a DTA with India through Broad.

3.6. Will you collaborate with any researchers not affiliated with Harvard to carry out this study?

☐ **No** ☒ **Yes:** *If yes, list which institutions they are affiliated with. If they are not affiliated with an institution, indicate that here. If yes, also indicate their responsibilities and scope of work in conducting this research.*

As this is an international study, below are our collaborators' affiliations and their scope of work:

Japan

Recruitment at the four Japan sites will be led by Dr. Tadafumi Kato, Professor and Head of Department of Psychiatry and Behavioral Science at Juntendo University. Dr. Kato has over 20 years of experience in clinical research of bipolar disorder and related mental illness. He is involved in the establishment of treatment guidelines of bipolar disorder and depression by the Japanese Society of Mood Disorders. He is also the chair of the Ethics Committee of RIKEN Brain Science Institute. Dr. Kato will oversee the recruitment in Japan and be in regular communication with the Broad Institute.

India

The India collection is overseen by Dr. Biju Viswanath, Associate Professor of Department of Psychiatry at National Institute of Mental Health and Neuro Sciences (NIMHANS). He will be the key leader for Indian cohort, overseeing and coordinating participants recruitment across 12 sites in India. He will also be in regular contact with the Broad Institute about recruitment updates.

The genetics data production will be led by Dr. Vijayalakshmi Ravindranath, Professor of Center for Neuroscience, Indian Institute of Science. She is also the Founder Director of Center for Brain Research at Indian Institute of Science. Dr. Ravindranath has expertise in research of pathogenic mechanisms in neurodegeneration and identification of drug targets.

Korea

Dr. Heon-Jeong Lee and Dr. Ji Hyun Baek will be the lead for Korea recruitment. Dr. Lee is the Professor of Psychiatry at Korea University College of Medicine. He is a clinical psychiatrist with particular expertise in bipolar disorders, mood disorders of early adulthood, chronobiology and sleep medicine. Dr. Lee has steadily devoted his effort to the genetic research on various phenotypes of mental illnesses over the two decades. Dr. Ji Hyun Baek is the Assistant Professor of the Department of Psychiatry, Samsung medical center. Her research interests are concentrated on finding genetic factors associated with bipolar disorders. Dr. Lee and Dr. Baek together will lead a total of 12 hospitals across Korea for participants recruitment. The sites are: Korea University Hospitals, Seoul National University Hospital, Budang Seoul National University Hospital, Samsung Seoul Hospital, Severance Hospital, Kyungbul National University Hospital, Busan National Univesity Hospital, Chunnam National University Hospital, National Center of Mental Health, Kyungsang

University Hospital, Eulji Hosptal and Ilsan Paik Hospital. They will oversee genomic data generation in Korea. In addition, they will join monthly project updates calls with the Broad Institute, and prepare progress reports.

Singapore

Singapore recruitment is based at the Institute of Mental Health in Singapore and will be led by Dr. Jimmy Lee. Dr. Lee has over 10 years' experience in psychiatric research, and has led a local study recruiting 2000 schizophrenia patients and 1000 healthy controls. He's interested in the clinical phenotyping and subsequent transdiagnostic evaluation of symptom domains between bipolar disorder and other serious mental illnesses. Dr. Lee will oversee data collection and training in Singapore, as well as communicating with the Broad PIs on updates.

Taiwan

Taiwan recruitment is overseen by Dr. Po-Hsiu Kuo, Professor and deputy director of the Institute of Epidemiology and Preventive Medicine, National Taiwan University. Dr. Kuo's main research focus includes genetic mapping and genomic studies for complex traits, psychiatric epidemiology, and bioinformatics studies, which fully supports this proposal. She has led many clinical studies for mood disorders in Taiwan, which successfully recruited more than 5000 individuals. In our project, Dr. Kuo and her team will coordinate the recruitment and training at all sites in Taiwan, which takes place in regional and central hospitals in the northern, central, southern and eastern of Taiwan. She will be the key leader communicating with the other two investigators about study updates.

3.7. Will your collaborators interact with human subjects, have access to identifiable data/specimens, and/or be responsible for the design, conduct, oversight, or reporting of the research?

☐ **No** ☒ **Yes:** *If yes, indicate if the collaborators will obtain their own IRB review.*

Collaborators will obtain the local institution's IRB approval prior to starting recruitment.

3.8. Will any institution conducting research activities as part of this study, including collaborators, rely on Harvard LMA for IRB review?

☐**No** ☒**Yes:** *If yes, list each relying institutions, their site responsible Investigator, and describe what research activities will be conducted there.*

The HSPH IRB has agreed to be the IRB of Record for this study under the auspices of the MOU with the Broad Institute.

Broad Institute: Broad will be involved in genomic data generation and data analysis.

HMS-MGH: Hailiang Huang is the PI of the study. His primary affiliation is MGH. They have confirmed they are not engaged in this research (See documentation attached to ESTR record). Dr. Huang is also Assistant Professor of Medicine at Harvard Medical School. He will oversee the study conduct, including data collection in Asian sites, data generation and analysis.

Virginia Commonwealth University also relies on HSPH IRB for review. The approval is uploaded in the ESTR.

4. Study Team

4.1. Describe the scope of work of the Harvard PI and research team. Indicate who is responsible for the design, conduct, implementation, and/or reporting of the research. Indicate who is responsible for the creation, design, and/or implementation of the study documents/tools.

Hailiang Huang is the lead PI for the study. He will oversee the study design, conduct and implementation throughout the project. Dr. Huang will communicate with site PIs via email and regular conference calls throughout the duration of the study to share updates and information from each of the recruitment sites, phenotypic collection and harmonization efforts, sample intake, data generation, data analysis and data sharing.

Jacab Taylor is responsible for creation, design, and training site research staff for the study questionnaires. For this grant he will work with Dr. Kenneth Kendler to develop the phenotyping battery, establish procedures for training those who will administer the phenotyping battery, and will coordinate the phenotypic quality control efforts. In addition, he will work with Dr. Kendler to help lead analytic efforts aimed at identifying genetic differences between groups of patients with different putative subtypes of bipolar disorder.

Zhenglin Guo, the project manager on the study, will oversee the study conduct and implementation under the supervision of PI. She will be responsible for routinely checking in with sites on study progress and reporting of the study updates to IRB or other regulatory committees.

Our Program Officer at the National Institute of Mental Health (NIMH), under whom this study is being funded, recommended that we add five *optional* questionnaires to the participant recruitment process.
There is one questionnaire related to the COVID-19 pandemic, filename: "A-BIG-NET COVID Assessment.docx".
There are four other questionnaires included to better capture study data and cultural norms. These are all new and optional questionnaires rather than revisions of tools included in previous versions of this submission. The four other questionnaire filenames include: "A-BIG-NET questionnaires_V2.0" (includes demographic and environment information) questions from which have been incorporated into the main *"A-BIG-NET_questionnaires.pdf*" file.

Also included now are "WHODAS.clean.pdf"; "PHQ-9_English.pdf" (Patient Health Questionnaire); and "GAD-7_English.pdf".

4.2. Describe the Principal Investigator's experience conducting research at the study site(s) and familiarity with the local research context.

Hailiang Huang Ph.D. is an Instructor in the Analytic and Translational Genetics Unit at Massachusetts General Hospital and the Director of Stanley Center Asia Initiatives at the Broad Institute of MIT and Harvard. Dr. Huang is also Assistant Professor of Medicine at Harvard Medical School. He has enduring interests and broad background in developing and using computational approaches to understand the relation between genomic variations and human diseases. He developed GWiS, a gene-based association test that has been used in many consortia to find genes associated with human complex disorders. He is a member of the International Inflammatory Bowel Diseases Genetics Consortium (IIBDGC) and has led its recent fine-mapping effort to resolve known genetic associations to variants with high causal probabilities. He is also leading a workgroup in the Psychiatric Genomics Consortium (PGC) to build a large-scale Asian schizophrenia cohort and use this cohort to understand the genetic architecture of schizophrenia in the Asian populations. In addition to his research activities, he has been organizing and teaching an annual statistical genetics workshop series in Shanghai (2015), Beijing (2017) and Taipei (2015 and 2016). These workshops have been attended by over 250 participants from more than 50 institutions, and helped to build global genetics research capacity.

Jacob Talor MD. is an instructor in the Department of Psychiatry at Brigham and Women's Hospital (BWH). His research interests include using genetic data to understand the biological basis of psychiatric disorders and the phenotypic features that index distinct forms of biologic risk within and across diagnostic categories. He led an international project to more precisely define the role of de novo genetic mutations in driving the known association between advanced paternal age and risk for several psychiatric disorders. Along with Dr. Robinson, and with guidance and mentorship from Drs. Hyman and Kenneth Kendler (a member of the Stanley Center's Scientific Advisory Board) he is also helping to lead the Stanley Center's Patient Stratification Initiative. The goal of this initiative is to integrate phenotypic and genetic variation among psychiatric patients in order to be able to more effectively stratify patients for further genetic, prognostic and therapeutic investigations

Tadafumi Kato M.D., Ph.D. is a professor in the Department of Psychiatry and Behavioral Science in Juntendo University. He has worked in the clinical department of psychiatry and focused on clinical research on bipolar disorder and related mental disorders using magnetic resonance spectroscopy and molecular genetics. In the last 20 years, he has focused on basic research on bipolar disorder, such as genetics, genomics, animal models, and so on. In the last decade, Dr. Kato has been organizing a trio-exome sequencing of bipolar disorder. He established a system to recruit patients with bipolar disorder by organizing a "Bipolar Disorder Research Network Japan (BDRNJ)". Through the advertisement by mail magazines and the web site, patients were enrolled. Informed consent and clinical assessment were done

through telephone interviews. Through this system, we recruited more than 230 families, enabling one of the largest collections of trio families of bipolar disorder.

Biju Viswanath M.D., Ph.D. is an Associate Professor at National Institute of Mental Health and Neuro Sciences (NIMHANS) in India. Dr. Viswanath's primary area of research includes understanding the molecular basis of etiology and treatment response in neuropsychiatric disorders. His training as a physician specializing in psychiatry and as a basic researcher in human stem cells enables me with a unique skill set to develop and execute translational psychiatry research. He's currently funded to examine genetic and cellular correlates of lithium response in bipolar disorder. He is also a key investigator in the Accelerator program for Discovery in Brain disorders using Stem cells (ADBS) program (https://ncbs.res.in/adbs/home) funded by Department of Biotechnology which aims to build a biorepository of human induced pluripotent stem cells for serious mental illnesses in India. The proposed study is in line with his interests in the genetic basis of bipolar disorder. He has established a network of investigators in India (https://www.med.unc.edu/pgc/about-us/people/pgc-ic-magic/) which has the capability to become an important contributor in the field of psychiatric genetics.

Heon-Jeong Lee M.D., Ph.D. is a professor of Psychiatry at Korea University College of Medicine. He is a clinical psychiatrist with broad experiences in treating patients with severe, treatment-resistant mental illnesses at the university-affiliated hospital. He has particular clinical expertise in bipolar disorders, mood disorders of early adulthood, chronobiology and sleep medicine. On the other hand, as an active researcher, Dr. Lee has steadily devoted his effort to the genetic research on various phenotypes of mental illnesses over the two decades. Since 2015, he has led the cohort named the Mood Disorder Cohort Research Consortium (MDCRC) as the principal investigator. The MDCRC, funded by the Korean government, is a multi-site naturalistic observational cohort of patients with early onset mood disorders, which will provide valuable multidimensional longitudinal data to clarify the key features and intervention target during the early phase of mood disorders.

Jimmy Lee M.D. is a practicing psychiatrist, clinician scientist and a Lead PI of the Phenomics Core at the Institute of Mental Health (IMH) Singapore. He has been engaged in psychiatric research for more than 10 years and has held several research grants ranging from clinical phenotyping to neuroimaging and biomarker studies. Notably, Dr. Lee had experience in studies related to the present submission on clinical phenotyping and genetics. He was involved in local studies on genetics in schizophrenia, recruiting 2000 individuals with schizophrenia and 1000 healthy controls. He has more than 100 peer-reviewed publications, co-authored 3 book chapters in psychiatry with a H-index of 34. He would also be especially interested in the clinical phenotyping and subsequent transdiagnostic evaluation of symptom domains between Bipolar Disorders and other serious mental illnesses.

4.3. Describe how the Principal Investigator will ensure that sufficient time is devoted to conducting and completing the research.

The PIs, Key Personnel and study staff from the Broad Institute, Virginia Commonwealth University, Johns Hopkins University and the National University of Taiwan will conduct monthly steering committee calls once the award period begins, and communicate via email throughout the duration of the study to share updates and information from each of the recruitment sites, phenotypic collection and harmonization efforts, sample intake, data generation, data analysis and data sharing. If site PIs are unavailable, they will ensure that a representative is on the call to participate in their place. Once we enter into the analysis phase of the project, we will schedule bi-weekly phenotype and analysis calls, in addition to the steering committee calls.

Once the project starts recruitment, Dr. Huang and Dr. Taylor will join the weekly calls with PIs and research staff from each study site addressing issues arising in participants recruitment and phenotyping collections. PIs will also communicate with study staff through emails on a daily basis for issues related to data collection.

In addition to conference calls, Dr. Huang or Dr. Taylor will visit 5 study sites per year to guarantee the progress of the project, as well as discuss data analysis and refresh training.

4.4. Describe how all research staff members are trained to ensure that they are adequately informed about the protocol and study-related duties.

We will have a study initial training for each site PI and their senior staff on the screening process, inclusion/exclusion criteria, and administration of the SCID5, MINI and these novel tools. Each site will, in turn, conduct a standardized training for all interviewers with brief appearances via video conferencing by PI Dr. Kendler (VCU) and co-I Taylor to introduce the study, discuss the overall strategy and answer specific questions that emerge.
All interviewers will have to pass an online test to demonstrate adequate understanding of each item on all data collection tools prior to beginning data collection.

Co-I Taylor and PI Kendler will also provide ongoing training to ensure uniformity in phenotypic measurements across sites. A document will be prepared and circulated containing detailed instructions for each item across all phenotyping instruments (a 'Q by Q'). co-I Taylor and/or the to-be-hired Clinical Psychologist working with PI Kendler will communicate regularly with interviewers across all sites in order to discuss any situations in which it was not clear how a particular item should have been scored. Based on these discussions, the Q by Q document will be updated regularly. Additions and other changes to the Q by Q will be documented in regular memos to all participating sites and there will be cross-site virtual meetings at least monthly to discuss these changes in detail. Throughout data collection, there will be regular communication with all clinicians administering tests about how to resolve ambiguities including at least monthly cross-site virtual meetings with participation from the US sites to discuss ambiguities in more depth and ensure uniformity in scoring across sites.

Besides study related training, all study staff are required to complete human subjects protection training before working on the project. There will be an initial training with site PIs on study procedures and required documentation. The project manager will be checking in with sites regularly to ensure study compliance. All local study staff will be trained on their country-specific mandatory reporting laws.

4.5. Describe the minimum qualifications for each research role (e.g., RN, social worker, phlebotomist, statistician), their experience in conducting research, and their knowledge of the local research context.

Project Managers on this study will be responsible for ongoing communications with collaborators at study sites to facilitate the study set up and genomic processing of incoming samples. Once samples are collected and DNA extracted at study sites, our Project Management team will steward them through the Broad's established sample submission process all the way through to data generation and manage the phenotypic metadata collection. The Project Managers (minimum MS in a biological or health sciences field of study) keep all members of the study team, both internal and external, well informed on project status as batches of samples are processed.

Broad Genomics Lab members (minimum BS in a biological field of study) respond to requests from Project Managers to send Broad barcoded matrix kits to our study sites for collection of DNA samples that are then shipped to the Broad. Once samples arrive to Broad Genomics, staff there quantify them in the samples lab and enter them into the genotyping pipeline for processing on GSA-MD arrays from Illumina. After genotyping data are generated, they are handed off to the Project Managers and Data Analysts for downstream analysis.

Computational biologists perform quality control to removes sample and genotype widely agreed in the field to be unreliable, and confirms sample sex, familial relationships, and general ancestry from the genotype data. Computational biologists have a minimum B.A. or B.S. with a major in biological and/or computational field of study.

Data analysts oversee the QC activities and perform association analysis including primary SNP association, meta-analysis of multiple datasets, and follow-up analyses examining heritability, polygenic risk, and pathway enrichment. All association analyses are performed by a postdoctoral associate in the lab (minimum Ph.D. with a focus on statistical and/or computational field of study) and overseen by Principal Investigator.

Psychologists at study sites will perform clinical interviews necessary for the diagnosis of the patients and the clinical assessment of the controls. Informed consent with participants is also performed by the psychologists. Psychologists have a minimum Bachelor degree with a major in psychology and are board certified in the country.

Clinical coordinators/Research assistants will participate in the recruitment of patients and the clinical assessment of patients' symptoms and courses. Over the course of the study, they will

help in collecting clinical and genetic data on the project, as well as performing DNA extraction in the lab. They will work under the direct supervision of the site PIs. The clinical coordinators/research assistants have a minimum of B.S. with a major in biology or medicine related field of study.

5. Study Design

5.1. Describe the study design type.

This is a case-control study recruiting 19,500 bipolar disorder cases and 15,000 controls in people of Asian descent. Genetic analysis will be performed to identify genetic variants associated with BP-I and explore phenotypic variance in BP-I and its relationship with genetic signatures and environmental risk factors. If funded, new study recruitment will take place in five countries in Asia, which covers 40% of the Asian population. Sites were selected on the basis of the following criteria: 1) proven track record of psychiatric research; 2) availability of research personnel and the necessary research infrastructure to be able to recruit thousands of participants; 3) existing trusted relationships from prior collaborations. Each of the countries where participants will be recruited from has enormous geographic and genetic diversity within and between them (1-3), which is likely to improve the ability of this research to answer the study objectives.

5.2. Does the study involve more than one participant group?

☐ **No** ☒ **Yes:** *If yes, identify each group here and throughout all applicable sections.*
Yes, the participant groups involve cases and controls.

5.3. Indicate the total duration of a participant's involvement.

The total study visit will be around 3 hours in total. It includes:

- Consenting: 30 minutes
- Questionnaires and assessments: 90-120 minutes
- Blood draw: 10 minutes

5.4. Indicate the total number of participants to be screened (if applicable) and/or enrolled (i.e., signed consent form). If the proposed research involves secondary data analyses only, indicate the number of data, documents, records, and/or specimens that will be obtained.

The four-year target enrollment of prospective collections across all of the sites is ~34,500. We expect to recruit an estimated 19,500 cases and 15,000 controls. This amounts to an average annual enrollment of more than 4,475 cases per year and almost 3,425 controls per year across all five countries. Below is a table indicating the planned recruitment numbers by each participating site:

Recruitment Milestones

Site	Type of Subject (Cases vs. Controls)	End of Y1 (cumulative) 9/1/22 - 8/31/23	End of Y2 (cumulative) 9/1/23 - 8/31/24	End of Y3 (cumulative) 9/1/24 - 8/31/25	End of Y4 (cumulative) 9/1/25 - 8/31/26	Total
Japan	Cases	350	950	1550	2000	3000
	Controls	200	500	800	1000	
India	Cases	2000	5000	8000	10000	20000
	Controls	2000	5000	8000	10000	
Korea	Cases	600	2000	3400	4000	6000
	Controls	400	1000	1600	2000	
Singapore	Cases	100	250	400	500	1000
	Controls	100	250	400	500	
Taiwan	Cases	500	1500	2500	3000	4500
	Controls	250	750	1250	1500	
Total	**Cases**	3550	9700	15850	19500	34500
	Controls	2950	7500	12050	15000	

5.5. List inclusion and exclusion criteria, including age ranges of all participants, and describe the screening process. Provide a rationale for any specific exclusion criteria.

All the sites from five countries use the same recruitment protocol to ensure study consistency. Potential cases will be recruited from outpatient and inpatient clinics with the following inclusion criteria:

1. Any subjects who meet DSM-5 the diagnostic criteria for Bipolar I Disorder;
2. Age at recruitment is 18-75 years old
3. Age of the first episode (either major depressive episode or manic episode) prior to age 50;
4. At least three years since the initial diagnosis of Bipolar Disorder.

Cases will be excluded if:

1. Persons who are currently involuntarily detained
2. Persons who are currently acutely psychiatrically illness
3. Anyone whose treating physician believes that participation could be harmful to them
4. Any history of a psychiatric illness caused by a neurologic condition (including mental disorders caused by epilepsy, cerebral tumor, CNS infections, or head trauma)

5. Any history of moderate or severe intellectual disability
6. Any history of a diagnosis of a mental disorder secondary drug or alcohol use

Selected controls will be recruited at each site and matched on ascertainment location, ethnicity, age at ascertainment and sex. Controls will be recruited following the inclusion criteria below:

1. Any subjects who screens negative for bipolar disorder I or II, schizoaffective disorder and schizophrenia
2. Age at recruitment is 18-75 years old

Exclusion criteria of controls are:

1. Any history of mania or psychosis or diagnosis of any DSM-5 bipolar diagnosis or psychotic disorder
2. Any history of a psychiatric illness caused by a neurologic condition (including mental disorders caused by epilepsy, cerebral tumor, CNS infections, or head trauma)
3. Any history of moderate or severe intellectual disability
4. Any history of a diagnosis of a mental disorder secondary to drug or alcohol use.

Note that a diagnosis of major depressive disorder (without psychosis) or a diagnosis of a substance use disorder itself will not be exclusion criteria for controls in order to avoid biases in our genetic findings with signals related to these relatively common conditions.

Rationales for exclusion criteria in cases:

1. We excluded the patients who are involuntarily detained in the hospital to make sure there's no coercion in participating the study
2. For the safety and wellness of patients, those who are under acute illness will be excluded from the study
3. Premorbid mental disorders and previous history of drug use are excluded to eliminate the confounding factors in the analysis.

5.6. Describe study procedures.

Participants will 1) undergo an interview with research assistants about basic demographic information, including information on putative environmental risk factors for bipolar disorder, and subject reports on past psychiatric diagnoses as well as history of affective and psychotic disorders in first degree relatives. 2) provide a small blood specimen (under 50 milliliters).

During the screening, research assistant will review medical charts / ask participants the year and month of the initial diagnosis to confirm the eligibility. Another study staff will double

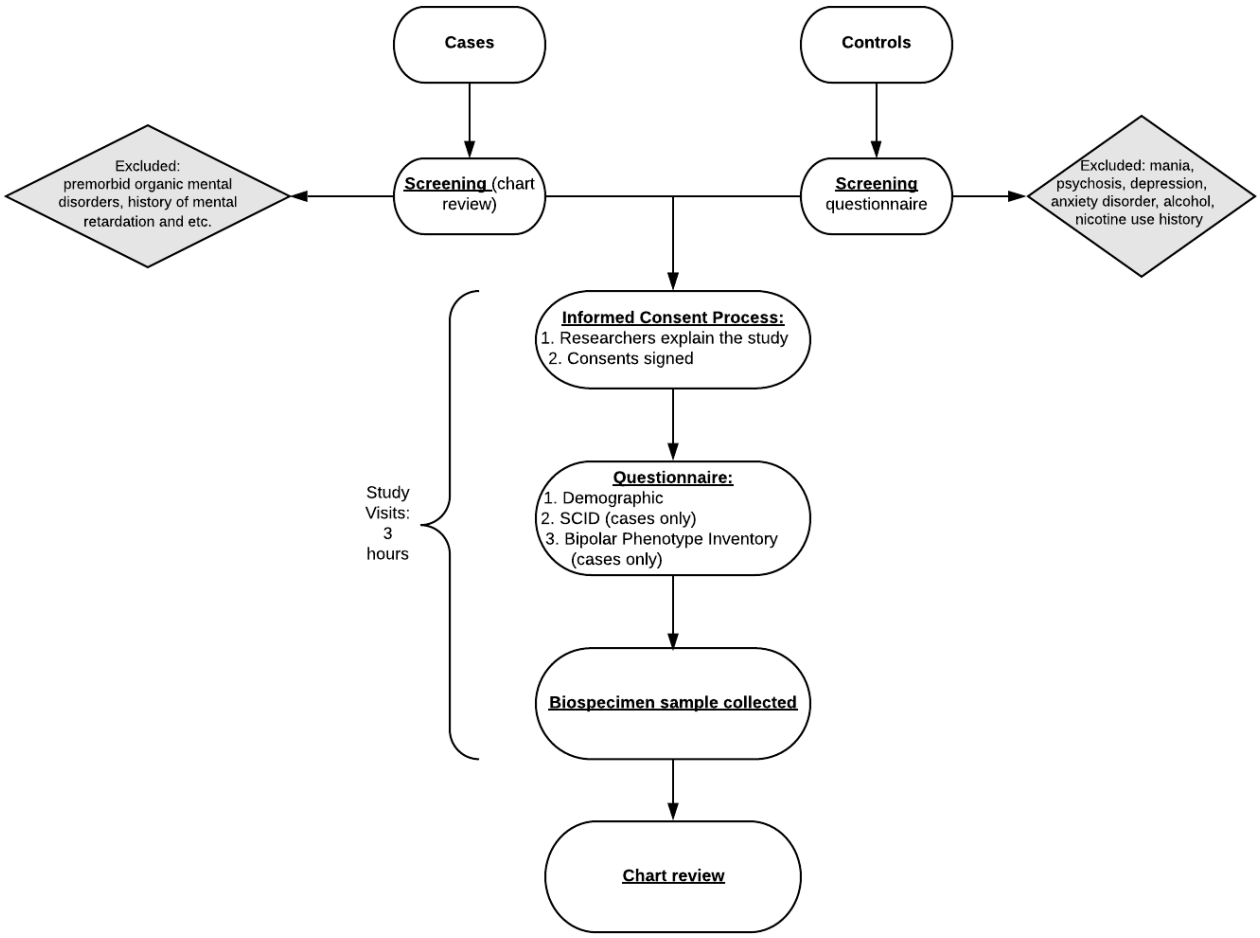
check the participant's eligibility to make sure all the criteria are met before enrolling to the study.

Additionally, cases will be interviewed by a clinician: 1) The mood and psychotic disorder modules from the Structured Clinical Interview for DSM 5 (SCID-5) will be administered to confirm that cases meet inclusion criteria and to collect item-level information about specific symptoms and course of illness and, 2) a custom brief instrument (the Bipolar Phenotype Inventory), in order to collect data on domains of functioning, psychopathology and other history. Items on this instrument will be specifically chosen to capture variation where there is existing evidence that such variation tracks with genetic heterogeneity. This instrument will be administered through patient interview, and, where available, informant interview and/or review of records.

The phenotyping battery is the whole questionnaire that will be administered by local researchers, including SCID and demographic information.

As stated, screening questionnaires, interviews, and blood tubes will be identified only with a subject ID number. The coding key and consent forms, plus any re-contact information, must and will be maintained in a coded spreadsheet in a secure location at each local site. Only authorized and trained research staff will have access to these files.

In addition, due to COVID-19 outbreak, all the participants coming to the clinic will be required to wear surgical masks. There will also be free masks at the front desk for participants to wear. Participants will be asked to take temperature, report if they have cough, shortness of breath, or other flu symptoms, travel and contact history. All the research staff will were PPE when conducting those study procedures.



5.7. Does the study involve the use of deception and/or incomplete disclosure?

☒ **No** ☐ **Yes:** *If yes, explain the use of deception/incomplete disclosure and describe why it is necessary to achieve the goals of the study.*

5.8. When all research-related study procedures are complete, are there plans for long-term follow up?

☒**No** ☐**Yes:** *If yes, indicate what data will be collected during this period*

The study doesn't have plans for long-term follow up. If the study have unique genetic findings, we may ask certain participants with or without those genetic variants if they are willing to participate in another study in future. The option to be contacted in the future study is indicated in the informed consent.

5.9. Does the study involve the collection of specimens (e.g. blood, cells, tissues, fluids, secretions, recombinant or synthetic nucleic acids, biological toxins, bacteria, virus, fungi, etc.)

☐ **No** ☒ **Yes:** *If yes, indicate the COMS Registration Number or plans to obtain COMS approval.*
The study will collect blood samples from participants.

5.10. Does the study involve the use of existing data, documents, records, and/or specimens for secondary analysis?
☐ **No** ☒ **Yes:** *If yes, indicate how, when, where, and from whom data, documents, records,*
The study involves review of medical records during screening and data collection process. The medical records will only be available at local research sites and will only be reviewed by researchers who have access.

5.11. Are there provisions for medical and/or psychological support resources available to participants (e.g., in the event of incidental findings, research-related stress)?
☐ **No** ☒ **Yes:** *If yes, describe the provisions and their availability.*

The findings from the systematic psychiatric evaluation may indicate information on accurate psychiatric diagnosis, presence or absence of psychiatric symptoms, and longitudinal course of psychiatric disorder, which will be provided to participants to manage overall psychiatric illness. Also, we will notify participants if their data assessed through the participation has a significant impact on health or requires immediate action for psychiatric treatment.

During the study process, if a participant feels stress or uncomfortable, the participant may take a break or skip the session. The study team will provide psychological support for the research-related stress. If a subject indicates potential harm to themselves or others, the study team will refer the subject to the specialist for consultation.

We do not currently plan to release results of genotyping analyses, including any incidental findings, to subjects, their families, or to third parties. This has also been our policy on previous genetic studies of similar disorders. The reason for this policy is that we do not believe that these results will have a straightforward or predictive interpretation, at an individual level, in the foreseeable future. We would only modify this policy in accordance with IRB standards, if at some point we are able to provide accurate and quantifiable risks to individuals for the disorders that we are studying. Because the clinical relevance of the data to be acquired is yet to be determined, no other individual information will be shared with study participants.

5.12. Describe the data and safety monitoring plan for the study. This plan should outline how study progress will be monitored throughout the lifecycle of the research to ensure the safety of subjects, as well as the integrity and confidentiality of data.

Monitoring Plan

The study will be overseen by Broad PI Hailiang Huang. Under Dr. Huang's supervision, the Broad project manager will routinely communicate and monitor the study progress:

1. Once the study is approved by both IRB, we will invite researchers from each site to Broad for initiation training (we may switch to virtual training due to pandemic). The training will cover:
    a. Screening and assessing participants with study tools and questionnaires
    b. Obtaining and documenting informed consent
    c. Use of logs, including enrollment log, staff delegation log and training log
    d. Adverse events report and Notes to File
    e. Collecting data with REDCap (REDCap is a password protected data collection tool)

2. The project manager will hold virtual conference with each site every 2 weeks, checking in on compliance and data collection:
    a. Study enrollment log will be reviewed bi-weekly for monitoring enrollment and compliance
    b. REDCap data forms will be reviewed for any missing data or queries
    c. New research staff and any unexpected events will be reported and discussed during the meeting

3. PI and the project manager will go site visits every year, monitoring the study and providing refresher training.
    a. The Broad project manager will randomly select 10% enrollment documents for review, including informed consents, enrollment logs, Notes to Files, and other documentation.
    b. Regulatory documents including local IRB approval, site training logs, delegation logs will also be reviewed on sites.
    c. Data confidentiality will be monitored during the site visits to make sure sites have secured places to store participants' documents.

Data Safety Plan

Throughout the study, confidentiality of participants will be the first priority. Plans to maintain confidentiality include the following:

1. All study staff across all sites will be trained, with ongoing supervision, to make confidentiality the first priority. All staff needing to access confidential health information or needing to interact with participants will complete Protection of Human Research training before being given authorization to do so.

2. All interviews will be conducted privately. If informant interviews are conducted, information received from individual family members will not be shared with any other family member.

3. All identifying information will be kept in locked areas and will not be shared with other investigators.
4. Consent forms and re-contact information, if given, will be stored in locked cabinets separate from demographic and diagnostic data.
5. The most serious risk would be identification of individuals in the publicly shared database. To prevent this, computerized data files provided to other investigators will not include names, addresses, date of birth, physical descriptions, or other individual-specific information. Only numerical ratings, clinical information, and final diagnoses will be provided. Published material will not identify subjects.
6. Some authorities on genetic studies have expressed concerns that it might someday become possible to identify subjects by comparing DNA test information from research with DNA information that might someday be found in medical records. This information could conceivably be used to deny insurance or employment to persons at risk for specific diseases. At present, there are no DNA marker data in routine medical records that could be used to identify subjects, and DNA information from research is not publicly available.

5.13. Are there any anticipated circumstances under which participants will be withdrawn from the research without their consent?
☒ **No** ☐ **Yes:** *If yes, describe the circumstances for withdrawal as well any associated procedures to ensure orderly termination, appropriate referrals, and/or follow-up care.*

6. Recruitment Methods ☐ N/A. *Skip to next section.*

Upload recruitment materials to the "Local-Site Documents" page in ESTR.

6.1. Indicate how, when, where, and by whom participants will be recruited. Provide a list of materials used to recruit participants, e.g., emails, posters, and/or scripts here.

Please note that sites recruitment materials are being reviewed by local IRB, we will submit the approved materials via an amendment. Neither consents nor recruitment materials will be implemented until both Harvard and the local IRB have approved.

Please note that only Singapore, Taiwan and Korea will start enrollment in this approval. Study details in Japan and India will be submitted in a modification when grant is awarded.

Singapore Recruitment

The Institute of Mental Health of Singapore will recruit 1000 participants (500 cases, 500 controls). All participants are over 21 years old (In Singapore below 21 years old is minor).

As Singapore is a multi-ethnic country, the study population will consist of Chinese and South Asians. Research team is well trained who understands best about local culture, religion of different ethnic groups and social delicacies. Recruitment flyers containing brief information about the study, inclusion/exclusion criteria and study contact information are presented to participants in local languages, including English, Chinese and Tamil. Recruitment conversations are done in the potential participant's preferred language by a bachelor's level Research Assistant.

Ascertainment and Screening of Cases:
Patients are recruited from both inpatient and outpatient services. For inpatient recruitment, only patients from open wards and are in a stable status will be enrolled to the study. In outpatient clinics, study flyers are handed to patients in the waiting room. Clinicians will explain the study and then refer patients who are interested in the study to the research team. Research team will describe the study in local language, give participants informed consent to read and answer all the questions they may have. Consenting process involving the collection of blood samples, structured clinical interviews and assessments, allowing access to information contained in their medical record, and broad consent to allow sharing of coded samples/data with other researchers.

Inclusion and exclusion criteria for cases: consistent with the study criteria.

Ascertainment and Screening of Controls:
We will recruit 500 controls, matched to the cases on age, gender, race and educational level attained and geographical location by advertising widely through study flyers in communities. Controls will also be recruited from other research databases with controls willingness to be recontacted for other studies. Those interested in participating will be directed to study staff located on-site, or will contact study staff by phone or email. Study staff will describe the study and perform a brief eligibility screening questionnaire. Individuals passing this initial screen will be invited to enroll, either while at the clinic during their routine visit, or during another scheduled time. The informed consent process for control participants will involve consenting to blood sampling, screening assessment, and broad consent for sharing. Participants will be compensated for transportation costs.

Inclusion and exclusion criteria for controls: consistent with the study criteria.

Taiwan Recruitment

All recruitment conversations are done in the potential participant's preferred language by a research assistant or study nurse, either in Mandarin Chinese or Taiwanese language. We enroll Taiwanese Han Chinese with both genders in this study. We plan to recruit 3000 cases and 1500 controls in total over four years. Because Taiwan legal adult age is set at 20 y.o., therefore, only adult men and women between 20 and 70 will be recruited within this study. Recruitment materials include flyers and posters of both hard copy format and electronic file format for social media recruitment. Flyers are present in Mandarin Chinese and containing

brief information about the study, inclusion/exclusion criteria are shown in the following section. Contact and more detailed information will be explained to patients by research assistants or psychiatrists. The target number of participants was carefully calculated based on estimated caseloads of patients with the diagnoses of interest in each of the study sites following several discussions with clinicians who practice at the study sites. We estimate approximately 50% participation rate, because participation is relatively slight-impact. In addition, we will re-contact patients (500~) and controls (500~) who participated in previous studies to invite them to join the proposed project. Those who agree to participate will re-sign the informed consent.

Cases – Patients with bipolar I disorder, or schizoaffective disorder-mania type are recruited primarily from clinical settings from both inpatient and outpatient facilities in central and regional hospitals. Patients will be recruited in the Northern, Central, Southern, and East-coast cities. In out-patient departments, psychiatrists briefly introduce the study to eligible patients, if patients are willing to join the study, then 1) referral to the research assistants to introduce the details of the study and sign the informed consents. 2) leave the contact information, the research assistants will contact them within seven days. For psychiatric inpatients, patients will be referred to a study assistant by psychiatrists when their mood is relatively stable. The study assistant introduces the details of the study and signs an informed consent in the ward, the detailed interview will be after the day of discharge. Cases are recruited predominantly on clinics and wards from a network of university hospitals in the four main regions in Taiwan.

Inclusion and exclusion criteria for cases: consistent with the study criteria and we will additionally exclude cases with substance induced bipolar disorder, mental retardation, schizophrenia, dementia, and HIV infection. Aboriginal in Taiwan will be excluded since our main target population is Han Chinese.

Controls - We ascertain controls from persons who present for treatment of general medical conditions at general medical hospitals that draw from similar catchment areas to the psychiatric facilities, as well as community controls. Controls must screen free from any psychiatric disorders history. We also use flyers of both hard copy format and electronic file format for social media recruitment. Willingness participants from social media actively contact and leave the message to the study team, study assistants will contact them to introduce the details of study and sign the informed consent and interview questionnaires. The local study team is the first point of contact for the patient and will explain the study and enroll controls into the study.

Inclusion and exclusion criteria for controls: consistent with the study criteria. Aboriginal in Taiwan will be excluded since our main target are Han Chinese

Korea Recruitment

Cases: We will recruit 4000 bipolar I patients in clinical settings from both inpatient and outpatient facilities. No cases will be enrolled or interviewed when they are currently an inpatient and cases will only be approached on the day of discharge. All participants are over 18 years of age. Both males and females will be recruited in the study.
All recruitment conversations are done by psychiatrists. Recruitment materials include flyers and posters. Flyers containing brief information about the study, inclusion/exclusion criteria and study contact information are handed out by the research staff to potential participants. The local study team is the first point of contact for the patient and explains the study and enrolls patients into the study.

Controls: We ascertain controls from persons who present for treatment of general medical conditions at general medical hospitals within similar catchment areas to the psychiatric facilities where cases are recruited. These individuals will be screened according to the overall study inclusion/exclusion criteria.

7. Consent Process

Upload consent form(s) and debriefing materials, if applicable, to the "Local-Site Documents" page in ESTR.

7.1. Describe how the research team will invite participants to take part in the research and obtain consent to participate. If the research team will not obtain informed consent, provide justification for requesting a waiver or alteration of consent (and/or parental permission).

Informed consent will be obtained for all adult participants who are participating in the studies. When conducting the informed consent process with study participants, trained research staff will determine the individual's level of comprehension of the material presented, based on whether the subjects are able to paraphrase the information in the consent form, to comprehend it adequately, and to ask reasonable questions about their participation. Any participant who does not demonstrate adequate understanding of the purpose, procedures, risks, benefits, emergency contacts, and payment issues will not be allowed to participate. All participants will be given a verbal explanation of the study purpose, procedures, risks, benefits, and payment and reminded that their participation is voluntary. Subjects who wish to do so will be encouraged to speak with their family members, friends, and physicians and show the consent form to them before making their decision to participate. The procedures are straightforward and outlined clearly in the informed consent process. After reading the consent forms or having assent forms explained, participants will be asked to provide a statement of their understanding of the research, the research procedure, any risks or discomfort involved, possible benefits of the study, and their right to withdraw at any time. If a participant indicates inability to comprehend provided material or is judged to be unable to comprehend this material, that individual will not be permitted to participate in the research. Once the participant confirms comprehension of the consent form and all study procedures, and agrees to participate he/she will be asked to sign the consent form.

7.2. Describe how the research team will document the consent process (e.g., participant/researcher will both sign and date the consent document; participants will thumbprint the consent document; electronic consent will be obtained and associated with the participant's research record). If the research team will not obtain signature and date, provide justification for requesting a waiver or alteration of documentation of consent (and/or parental permission).

We will use the paper consent form. Participant and researcher will both sign and date on the consent document. For participants who are illiterate, participants will fingerprint the consent document and researchers will write printed names and date for them. The consent process will be documented on the study enrollment log.

7.3. Will participants be offered a copy of the consent information?
☒ **Yes** ☐ **No:** *If no, explain why not.*

7.4. If consent will be obtained in a language other than English, identify the language(s) that consent information will be provided, who will be responsible for translation, and the provisions for communicating this information to participants. ☐ **N/A**
All consent forms will be in their native languages. For example, consents in Japan will be in Japanese; consents in Singapore will be in English, Chinese and Tamil; consents in India will be in Kannada, Tamil, Hindi, Telugu, Malayalam depending on the cities; consents in Taiwan will be in Mandarin Chinese or Taiwanese. All consenting conversations will happen in participants' preferred language in local hospitals/institutions.

The translation will be done jointly between Broad Institute investigators and site investigators, and the consent form will be submitted for local IRB review.

7.5. If the research involves deception and/or incomplete disclosure, describe the debriefing process. Explain when participants will be debriefed, who will debrief them, and how they will be debriefed. ☒ **N/A**

7.6. If the research involves secondary use of existing data, documents, records, and/or specimens, and the research team will not obtain consent, describe how consent was originally obtained. Additionally, either upload the original consent form to the ESTR record or confirm that the original consent process obtained participants' permission to share or use their data/specimens for future research projects. ☒ **N/A**

8. HIPAA Privacy Protections ☒ N/A. *Skip to next section.*

HIPAA applies to US-based research involving the collection or use of protected health information (PHI) from a hospital, health center (including the Harvard Dental Center), health plan, or health insurance plan (i.e. a covered entity). The Privacy Rule will not directly regulate researchers who are engaged in research within organizations that are not covered entities even

though they may gather, generate, access, and share personal health information (PHI). The Privacy Rule applies only to individually identifiable health information held or maintained by a covered entity. Individually identifiable health information that is held by anyone other than a covered entity, including an independent researcher who is not a covered entity, is not protected by the Privacy Rule and may be used or disclosed without regard to the Privacy Rule.

8.1. Explain how the Privacy Rule applies to this specific research project.

8.2. Describe the covered entity involved in this research that holds or maintains the PHI that will be used by the researchers.

8.3. Describe plans for obtaining authorization to access protected health information or provide the rationale for a waiver of authorization.

9. Research Subject to the European Union (EU) General Data Protection Regulation (GDPR)☒ N/A. *Skip to next section.*
GDPR applies to research involving the collection of "personal data" from research subjects who are located in the EEA. This includes biospecimens. The EU/EEA includes the 28 states of the European Union (Austria, Belgium, Bulgaria, Croatia, Republic of Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, & United Kingdom) and four additional countries: Iceland, Liechtenstein, Norway and Switzerland.

2.1. Describe plans to collect and/or obtain "pseudonymized data" (e.g., coded data) and/or identifiable data and/or biospecimens from participants in the EEA.

3. Research Subject to the Family Educational Rights and Privacy Act (FERPA)
☒ **N/A.** *Skip to next section.*
FERPA applies to research involving the collection of individually identifiable information from student records or personal education information from an education program (defined as: any program principally engaged in the provision of education, including, but not limited to, early childhood education, elementary and secondary education, postsecondary education, special education, job training, career and technical education, and adult education).

3.1. Describe plans to collect and/or obtain individually identifiable information from student records or personal education information from an education program.

4. Vulnerable Populations ☐ N/A. *Skip to next section.*

4.1. Identify all vulnerable populations (e.g., children; pregnant women, human fetuses, neonates; prisoners; elderly; economically disadvantaged; employees or students of the

investigator or sponsor; undocumented individuals; refugees; racial and/or ethnic minorities; illiterate or low-literacy; military personnel; terminally ill; cognitively impaired or mentally ill; persons with a stigmatizing disease or condition, e.g. AIDS/HIV, etc.) and describe safeguards to protect their rights and welfare.

Pregnant women and children under 18 are not eligible to participate in this study.

5. Risks

Risks may be physical, psychological, social, legal, reputational, and/or financial.

5.1. Describe the reasonably foreseeable risks, discomforts, and/or inconveniences to participants and/or the group/community to which they may belong. Indicate the probability, magnitude, and duration of each risk.

Blood Draw/DNA/RNA Analysis

The risks of venipuncture include brief minor pain related to needle insertion and, more rarely, bruising or blood clotting at the catheter site, which typically resolves within a few hours.

Clinical Assessment

Subjects may become tired during the assessment, and/or might feel uncomfortable discussing personal topics during the clinical interview. The researchers may ask questions related to participants' childhood experience, adverse experiences, income, family history and psychiatric histories. Those questions may be sensitive and participants may feel uncomfortable or anxious during the interview.

Breach of the subjects' confidentiality

The data we are collecting includes participants' personal information. The breach of confidentiality could lead to the identification of the participant and their employer, insurance company or others may have access to their health information. This could have damage to participants' employment, insurance and their social status.

5.2. Identify whether any of the information collected, if disclosed outside of the research, could reasonably place the participant at risk of criminal or civil liability or be damaging to the participant's financial standing, employability, insurability, or reputation.

One potential risk is a breach of the subjects' confidentiality, the names, and the IDs that could lead to the identification of the person. This could lead to their employer, insurance company, or others finding out that the subject participated in a research study. This could result in damaging to participant's employment, less opportunity to get insurance, and potential financial loss.

The disclosure of the participants' health information may lead to mental health stigma, especially in some rural areas. The stigma may lead to discrimination to participants, including lack of social support and fewer opportunities to work.

5.3. Outline provisions in place to minimize each risk identified above.

Blood Draw/DNA/RNA Analysis

The risks of venipuncture include brief minor pain related to needle insertion and, more rarely, bruising or blood clotting at the catheter site, which typically resolves within a few hours. All blood drawing equipment is disposable and is new from the package, and blood draws will be performed by an experienced phlebotomist. The entire chain of information related to blood sampling, from tubes containing blood, to stored DNA/RNA samples, to entering of information into a computer, will be identified by a random subject code number. No information that could be used to ascertain the identities of individual subjects will be present on any of the laboratory measures or records pertaining to the blood samples.

Clinical Assessment

Subjects may become tired during the assessment, and/or might feel uncomfortable discussing personal topics during the clinical interview. They may take a break or stop the interview at any time, and can decline to answer any question they do not feel comfortable answering. All information collected in this study will remain strictly confidential except where required by law.

Loss of Confidentiality

Protection against loss of confidentiality in the proposed project follows the same principles and procedures per local IRB. Strict standards of confidentiality are maintained. Subjects will remain anonymous in all publications. Data will be stored indefinitely for future research. There are potential liabilities to subjects if third parties (e.g. insurance companies or employers) discover that they are at increased risk for specific disorders. This risk is minimized by the confidentiality procedures described above.

To minimize the risk of any individual being identified by having clinical and genetic data publicly available, we plan to adhere to the NIH's policy for sharing of data as promulgated by the National Human Genome Research Institute (NHGRI). The identities of research subjects will not be disclosed to the data repository and all data will be submitted in a coded fashion. Data will be de-identified according to the following criteria: the identities of data subjects cannot be readily ascertained or otherwise associated with the data by the repository staff or secondary data users; the 18 identifiers enumerated at section 45 C.F.R. 164.514(b)(2) (the HIPAA Privacy Rule) are removed; and the submitting institution has no actual knowledge that the remaining information could be used alone or in combination with other information to identify the subject of the data. The data repository will never receive the code or any other information that would enable the identification of the individuals who are the source of the data. Identification of a specific individual through genetic variation data in the data repository

will require comparison with such data from another identifiable DNA sample from the same person. It is anticipated that technological and analytical capacity available to the public is likely to enhance the feasibility of such identification in the future. Similarly, the phenotype data deposited in the NIH data repository may include information about disease status and characteristics that are not individually identifiable; however, some characteristics may be shared in common among population subgroups.

To minimize risk of loss of confidentiality we will also report and adhere to our institutions' IRBs, verifying that the submission of data to the repository and its subsequent sharing to the scientific community are consistent with the informed consent obtained from participants. Also, access to the data through the repository is limited to those researchers with a relevant research plan. The procedure for gaining access to the data requires that the investigator and the institutional official certify that they will make no attempt to try to link data to individual subjects and that any potential for subject identification will be considered and defended against.

Protection due to COVID-19

Study will be conducted under the local hospital's policy and local health agency's guidance during the pandemic. In general:

1. Temperature will be taken before entering the clinic.
2. Anyone coming to the clinic is required to wear surgical masks. There will be free masks at the front desk for participants to wear before entering the clinic.
3. All healthcare workers are required to wear PPE when conducting study procedures.
4. Anyone who comes to the clinic has to go through screening questions, for example if they have cough, shortness of breath, fever, or other flu symptoms. Participants will also need to report recent travel and contact history.
5. If potential cases are identified, clinicians will report the case and refer to the special clinic for further COVID testing.

6. Benefits

6.1. Describe the potential benefits to individual participants, if any, and/or society. If there are no direct benefits, state that here. Note: payment/compensation is not a benefit.

There are no direct or immediate benefits to research participants.

If this study contributes to the identification of specific genes that contribute to individual susceptibility to bipolar disorder, then it may be possible in the future to develop better diagnostic tests, improved treatments, or preventive measures based on this knowledge. This might benefit persons in the future (possibly including some of these subjects or their descendants) who have or are at risk for bipolar disorder, schizoaffective bipolar type and related disorders. Risks involved in participation are minimal, particularly given the importance of the knowledge that may be gained from this study.

7. Participant Privacy

7.1. Describe provisions to protect participants' privacy (their ability to control and limit the extent, timing, and circumstances of sharing information about themselves with others, e.g., the use of a private interview room) and to minimize any sense of intrusiveness that may be caused by study questions or procedures.

Electronic data collection

All data collection instruments have electronic versions. A proprietary, SQL-based database (REDCap) has been established and will be used in the proposed study. REDCap is accessed using a password protected web-based client interface through which clinician interviewers directly enter participant responses. Access to specific content on REDCap is customizable and can be sharply limited; in our system, each collaborating site can only view data on the participants ascertained by them.
Screening questionnaires can be completed on ipad offline and, for eligible participants, uploaded later into the REDCap database. If no consent is obtained the screening data will be destroyed.

For cases, the interview will be conducted using ipads. All ipads are password protected and entry to the REDCap database also is password protected on several levels. Data entered by one interviewer is not viewable by another interview, even at the same site unless specific additional permissions are granted by the database administrator. Furthermore, as noted earlier, data collected at one collaborating site are not viewable at another site.

Documents with personal identifiers (i.e., signed consent forms, re-contact information, and the site coding key) are maintained at each site and stored separately from other study materials in locked file cabinets and/or password-protected databases. These source documents and databases are never transmitted between sites and personnel from one site are not able to contact a participant who was ascertained by a sister site.

All genetic data is maintained at the Broad Institute and collaborators at India and Korea (India and Korea are not allowed to ship biological samples overseas). Transmission of these data to NIMH databases proceeds according to our data sharing agreement. Our consent forms are consistent with the NIMH guidelines for submissions to the control-accessed database.

Participant confidentiality

Throughout the study, confidentiality of participants will be the first priority. Plans to maintain confidentiality include the following:

- All study staff across all sites will be trained, with ongoing supervision, to make confidentiality the first priority. All staff needing to access confidential health information or needing to interact with participants will complete Protection of Human Research training before being given authorization to do so.

- All interviews will be conducted privately. If informant interviews are conducted, information received from individual family members will not be shared with any other family member.
- All identifying information will be kept in locked areas and will not be shared with other investigators.
- Consent forms and re-contact information, if given, will be stored in locked cabinets separate from demographic and diagnostic data.
- The most serious risk would be identification of individuals in the publicly shared database. To prevent this, computerized data files provided to other investigators will not include names, addresses, date of birth, physical descriptions, or other individual-specific information. Only numerical ratings, clinical information, and final diagnoses will be provided. Published material will not identify subjects.
- Some authorities on genetic studies have expressed concerns that it might someday become possible to identify subjects by comparing DNA test information from research with DNA information that might someday be found in medical records. This information could conceivably be used to deny insurance or employment to persons at risk for specific diseases. At present, there are no DNA marker data in routine medical records that could be used to identify subjects, and DNA information from research is not publicly available.

8. Data Confidentiality

8.1. Indicate the identifiability of the data/specimens:

☐ Data/specimens will not contain any direct or indirect identifiers (anonymous data).
☐ Data/specimens will contain direct or indirect identifiers, but the research team will remove them upon receipt (de-identified data).
☒ Data/specimens will contain indirect identifiers (i.e., number, letter, symbol, or combination thereof) and the research team will maintain a key that links identifiers to individual participants (coded data).
☐ Data/specimens will contain direct identifiers (identifiable data).
☐ None of the above; describe:

8.2. Have any identifiable data/specimens been de-identified for use in this research study?
☒ No ☐ Yes: *If yes, describe how you will prevent any re-identification.*

8.3. Identify where data/specimens will be stored (e.g., on campus at Harvard or remotely, in a specimen laboratory) and describe the provisions to maintain confidentiality (e.g., password protection, encryption, locked filing cabinets, etc.). Refer to the Investigator Manual and the Harvard Research Data Security Policy for additional information.

Singapore
We will collect blood samples from both cases and control participants. About 10ml of blood will be collected by a phlebotomist or physician on the research team following the hospital standard protocols. The blood will be stored at a lab in the Institute of Mental Health in

Singapore for DNA,RNA, PBMC (peripheral blood mononuclear cells) and plasma extraction. The extracted DNA sufficient samples will be shipped to the Broad Institute of data generation.

Taiwan

We will collect blood samples from 2,400 cases and 1,200 control participants in Taiwan, for extraction of DNA. As part of this collection effort, participants are fully informed about all aspects of the study, including broad consent for data and sample sharing with other researchers, and written consent will be obtained. Approximately 25 ml of blood will be collected from each participant by a phlebotomist or physician on the research team or clinician of the recruiting hospital or clinic following protocols used in our previous funded projects. Samples will be centrifuged and stored frozen within 2 hours after blood draw for DNA isolation and QC of extracted DNA will be performed by OD and concentration record as well as gel electrophoresis to confirm integrity. Low pass whole-genome sequencing sufficient DNA will be extracted through methods abiding to either Phenol-Chloroform extraction method or spin column based methods (Qiagen, QIAamp DNA Blood Mini Kit). No additional basic blood tests will be performed prior to DNA extraction.

As an administrative update: Taiwan IRB approved Taiwan site to collect 2g of stool samples from participants to undergo next-generation sequencing technology to screen for all types of microorganisms for analysis of the intestinal microbiota in the human body. The stool samples will be stored and analyzed in Taiwan as Taiwan's own project. The Broad study team have requested Taiwan site to separate the stool samples collection in another consent form and will submit the updated version of consent once it's approved by local IRB.

In addition to the cohort ascertainment in Taiwan sites proposed in this project, DNA from the existing mood disorders cohort in Dr. Kuo's and collaborators' lab of approximately 1000 (BP cases and controls) and participants from our ongoing studies will also be invited for re-contacting the patients and re-consent to join this project. The protections for these samples are identical to the ones being ascertained in this proposal. All data will be coded and key linking data to living individuals is only accessible to study staff with access authorized by NTUH's IRB.

In addition, the existing DNA from the Taiwanese Schizophrenia trios cohort of nearly 1700 families (1700 patients with schizophrenia and 3400 unaffected relatives) will also be included under this proposal for analyses across mental disorders (LOS Glatt and Tsuang). Schizophrenia data from the phenotype interviews and the DNA can only be linked with a living individual through a random serial number as the unique ID. An encrypted database linking the participant ID and genetic ID is stored in-country and the local study team has the only additional code to link materials with individuals. Also, the participant's name and contact information is collected on the consent form, which is then stored in a locked cabinet and never leaves the country of origin.

Korea

10 ml of blood is extracted from the vein of the forearm, and the extracted blood is stored in the EDTA tube and frozen at -20 C. The collected samples will be shipped to the Korea University Hospital every three months and stored in the -70 C deep freezer. Sequencing will be performed in the specialized testing agency (Macrogene) in Seoul, South Korea. The remaining gDNA will be stored at the Korea University Hospital.

8.4. Indicate whether any data/specimens will be transferred/transmitted and describe the plan to share the data/specimens (e.g., outside of Harvard, to other researchers, to collaborators). Indicate who may request access and how. If data/specimens will be transferred/transmitted/shared, describe how, when, and to whom.

We will follow the NIH Genomic Data Sharing Policy and NIMH policies on sharing data. All subjects will be explicitly consented for future research use and broad data sharing. The phenotypic data collection will be harmonized with other major psychiatric genomics projects in ancestrally diverse populations to the extent possible, e.g., the Populations Underrepresented in Mental Illness Association Studies (PUMAS), to maximize its utility in future studies. We will deposit the anonymized genotype and phenotypic data to NIMH funded National Data Archive (NDA), once the quality control procedures are completed, which we expect to take place within three months after the final data production in early Year 5 of the project. The data will be released within six months after the data submission or at the acceptance of initial publication, whichever occurs first.

We will publicly release genomic summary results (GSR) generated from our analyses, including the allele frequencies, effect size estimates, p-values and etc, both through an interactive web portal and as downloadable files hosted at the NHGRI-EBI GWAS Catalog. We will release GSR as soon as possible but no later than the acceptance of initial publication.

We will also contribute all the data to the Psychiatric Genomics Consortium (PGC) Bipolar Workgroup and the Bipolar Sequencing Project (BSC) to further increase the power for psychiatric genomics research activities. We will make the contribution no later than the acceptance of the initial publication.

Biological samples from Taiwan, Singapore and Japan will be banked at the Broad Institute Genomics Platform; biological samples from South Korea and India will be banked at their respective facilities due to local regulations that either prevent them from leaving their country of origin, or prevent them from being banked in a foreign country. The study team will maintain a centralized sample manifest so that all samples can be located, tracked and be requested for future studies with proper legal and compliance approvals.

8.5. Indicate whether participants' permission will be obtained to share their data/specimens and/or use their data/specimens in other future research projects.
Data sharing and use of specimens are described in the consents. By signing the consent forms, participants will also agree to share their data which is described there.

8.6. Indicate who is responsible for data/specimen management and how the research team and/or other collaborators are permitted access to information.

PIs and the project manager are responsible for data and specimen management.

We will identify where the data will be available and how to access the data in any publications and presentations that we author or co-author using these data. As we will be using the NDA, which is an NIH-funded data staging platform, it has policies and procedures in place that will provide governed data access to qualified researchers, fully consistent with NIH data sharing policies and applicable laws and regulations. These data will be shared with investigators working under an institution with a Federal Wide Assurance (FWA) and could be used for secondary study purposes such as finding genes that contribute to the risk architecture of bipolar disorder. The names and Institutions of persons either given or denied access to the data, and the bases for such decisions, will be summarized in the annual progress report.

8.7. Indicate how long data/specimens will be stored and describe the plans at the end of the storage period (e.g., are data/specimens destroyed, returned to data/specimen provider, etc.).

Data will be stored indefinitely, as a resource for future research.

9. Data/Statistical Analyses Plan

9.1. Describe plans for analysis (including the statistical method, if applicable).

We will carry out a range of analyses to discover new genetic associations with BP-I across the allelic spectrum in East and South Asian populations, examine the comparative genetic architecture of BP-I across major world populations and with other major neuropsychiatric disorders, and perform a novel statistical fine-mapping analysis that leverages the multi-ancestry genomic diversity and pleiotropy across psychiatric disorders to identify putative causal variants. The goal is to explore the genetic "validity" of various BP-I subtypes and fit models with joint genetic and environmental risk factors. This proposal will dramatically increase the worldwide diversity of genetics data on BP, an important step to accelerate gene discovery in this disorder and advance global mental health equity.

We will complete the proposed study in five years (Figure below), performing the primary and follow up analyses in an iterative manner. Starting in year 2, we will develop analytic pipelines for both genetic and phenotype data. In year 3, we will perform a freeze for both genetic and phenotypic data so that the pipeline can be fully tested, polished and completed. After generation of the full data set early in year 5, we will use the developed genetic and phenotypic analysis pipelines to quickly complete the analyses and

write the relevant manuscripts.

Aim	Description	Year 1 1H	Year 1 2H	Year 2 1H	Year 2 2H	Year 3 1H	Year 3 2H	Year 4 1H	Year 4 2H	Year 5 1H	Year 5 2H
1	Training			R		R		R			
	Translation										
	App. dev.										
	Recruitment (%)		19	50		81		100			
	Sample per country (in hundreds), shown as #case/#control										
	Taiwan		4/2	8/4		8/4		4/2			
	Japan		3.5/2	6/3		6/3		4.5/2			
	S. Korea		5/2	10/3		10/3		5/2			
	Singapore		1/1	1.5/1.5		1.5/1.5		1/1			
	India		20/20	30/30		30/30		20/20			
2	Sequencing*	9,600		10,000		8,500		7,800		7,700	
	Quality control										
3	Primary analyses										
	Follow up analyses										

(Timeline of the proposed study. 1H: the first half of the year, 2H: the second half of the year. *Additional samples for sequencing available from a NIH funded study in Pakistan. R= Re-training to ensure phenotyping quality.)

9.2. Is there a sample size/power calculation?

☐ **No** ☒ **Yes:** *If yes, describe the calculation and the scientific rationale, and, if applicable, by site and key characteristics such as participant demographics.*

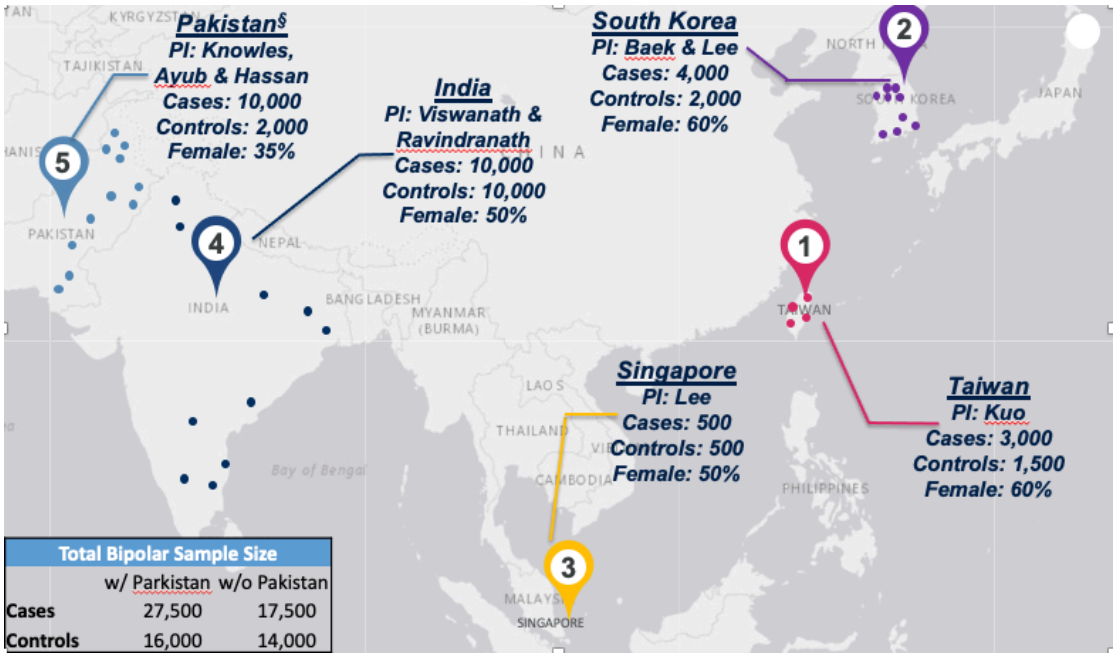
First, BP genetic studies are severely underpowered. Genetic discovery in BP lags behind other key complex disorders in the sample size and number of genetic associations identified. The Bipolar Sequencing Consortium (BSC) also lags behind the schizophrenia exome sequencing project (SCHEMA, manuscript forthcoming) in sample size/power to detect rare risk variants. SCHEMA has found ultra-rare coding variants that implicate 10 risk genes for schizophrenia using 24,000 cases and 97,000 controls. BSC, with 14,210 cases and 14,422 controls, has not identified significant associations with any individual gene but shows promising aggregate results: novel singleton protein truncating variants in genes intolerant to loss-of-function mutations are enriched in BP patients. Additional samples are needed to drive the discovery of both common and rare variants.

Second, most samples in BP genetic studies have minimal phenotyping information. Like other psychiatric disorders, BP is clinically heterogeneous. Psychiatric genetic studies built on minimal phenotyping have lower estimated heritability, and findings from such studies are less specific to the disorder of interest. Preliminary results from the PGC BP Working Group are consistent with this pattern, with lower heritability estimates from self-reported diagnoses (12%) versus clinical diagnoses (20%). Furthermore, even in clinically diagnosed samples of patients with BPI, phenotypic heterogeneity may reflect genetic heterogeneity. Heterogeneity may be particularly relevant to genetic discovery in BP. In contrast with schizophrenia, where more recent genetic studies have replicated 107 of the 108

previously reported loci, two of the 30 previously reported loci were not replicated in the most recent PGC BP study (manuscript forthcoming). **Third, most samples in the latest PGC BP GWAS were of European ancestry, exacerbating healthcare disparities by creating uncertainty in how applicable BP genetic findings may be to other populations.** Using only European cohorts will also miss important disease causing variants absent or rare in Europeans.

Recruitment Demographics

Below is a map featuring the six recruitment sites that we intend to collect from and plans for the study which including the site names, PI names, target sample size and proportion of females.



§ Existing DNA Funded by another grant

Please note: Japan will be supported under other funding so it's not included in the recruitment map to NIH grant.

10. Costs and Compensation ☐ **N/A.** *Skip to next section.*

10.1. Identify any costs that participants may incur during the study, including transportation costs, childcare, or other out-of-pocket expenses.

Participants will be recruited from their routine clinics visits, so there won't be any transportation costs. If additional visits are needed for the research, we will reimburse for the transportation cost.

10.2. Identify remuneration that participants may receive during the study. Specify the amount, timing of disbursement, and method (e.g. money, gift cards, in-kind, incentives,

raffles, and transportation). Describe how compensation will be calculated and paid if a participant withdraws. If any participant will receive a single payment more than $100, or $600 or more in one calendar year, refer to Harvard University Financial Policy on Human Subject Payments.

Taiwan

We offer the money (NT$500, about $17 US dollars) to the participants when they complete the questionnaire interview and draw blood. If participants withdraw from the study and they have already finished 1/3 to 1/2 of the interview, we will offer half of the compensation (NT$250, about $8.5 US dollars). On the other hand, participants have already finished more than 1/2 of the interview, we will offer all of the compensation (NT$500, about $17 US dollars) if they drop out of the study.

Singapore

We plan to give S$60 (about 44.5 USD) for a complete assessment, S$30 (about 22 USD) for incomplete assessments (participant withdraws or unable to finish assessment for any reason) and S$20 (about 15USD) for screen failure (on SCID/MINI). The compensation plan will also be described in the informed consent.

Korea

We will offer a set amount of compensation at the completion of the participation procedures. For the case group, they will receive KRW 70,000 (about 60 USD). For the participants in the control group, they will receive KRW 50,000 (about 43 USD). The compensation will be deposited into the participants' bank account within about one month.
Please note for sites in Japan and India, the detailed compensation plan is under review by local IRB. We will update this section after local IRB approval.

11. Sharing Study Results ☐ N/A. *Skip to next section.*

11.1. Describe the plan to share study results with individual participants, the participant group/community, and/or others.

Singapore

We do not currently plan to release results of genotyping analyses, including any incidental findings, to subjects, their families, or to third parties. This has also been our policy on previous genetic studies of similar disorders. The reason for this policy is that we do not believe that these results will have a straightforward or predictive interpretation, at an individual level, in the foreseeable future. We would only modify this policy, in accordance with IRB standards, if at some point we are able to provide accurate and quantifiable risks to individuals for the disorders that we are studying. Because the clinical relevance of the data to be acquired is yet to be determined, no other individual information will be shared with study participants.

Taiwan

We do not currently plan to release results of genotyping analyses, including any incidental findings, to subjects, their families, or to third parties. This has also been our policy on previous genetic studies of similar disorders. The reason for this policy is that we do not

believe that these results will have a straightforward or predictive interpretation, at an individual level, in the foreseeable future. We would only modify this policy, in accordance with IRB standards, if at some point we are able to provide accurate and quantifiable risks to individuals for the disorders that we are studying. Because the clinical relevance of the data to be acquired is yet to be determined, no other individual information will be shared with study participants.

Korea

We do not currently plan to release results of genotyping analyses, including any incidental findings, to subjects, their families, or to third parties. This has also been our policy on previous genetic studies of similar disorders. The reason for this policy is that we do not believe that these results will have a straightforward or predictive interpretation, at an individual level, in the foreseeable future. We would only modify this policy, in accordance with IRB standards, if at some point we are able to provide accurate and quantifiable risks to individuals for the disorders that we are studying. Because the clinical relevance of the data to be acquired is yet to be determined, no other individual information will be shared with study participants.

We will notify the participants if the data assessed through the participation has a significant impact on participants' health or requires immediate action for psychiatric treatment.

12. Research Related Injuries ☐ N/A. *Skip to next section.*

12.1. Describe plans for medical care and compensation for research-related injuries.

Participants will be subject to all safety precautions taken during the usual medical evaluation. Due to the nature of the research where only single time-point information is collected, the risk of an adverse event from current study procedures is minimal. However, appropriate protocols for dealing with adverse events will follow those developed for our previous studies. Access to licensed clinicians including treating physicians who are able to respond to such events will be facilitated by the presence of clinic and study staff.

13. Reportable Events

13.1. Outline plans for communicating reportable events to the IRB, Sponsor, or others as applicable (e.g., adverse events, unanticipated problems involving risks to participants or others, breach of confidentiality).

Adverse events and/or unanticipated problems involving risks to participants or others will be documented on an Adverse Event Form. Sites will report the unexpected events to the local IRB following their institution's policy. In addition, this information will be reported to the PI within 5 business days from the time the study team becomes aware of the adverse event and the local PI will inform Broad Institute within 5 days from the time that he becomes aware of the adverse event. The site PI will call a meeting to discuss the adverse event and will report it to the local IRB in accordance with institutional policies. In tandem, the site PI will report this information to agents of Harvard and the Broad within 5 business days. Documenting and reporting adverse events will be one of the items on which all staff are trained and will be reviewed when Broad Institute staff members visit the sites.

14. Regulatory Compliance

14.1. Describe plans for monitoring regulatory compliance. The monitoring plan should include how you will ensure proper record keeping, retention of required regulatory documents and participant files, and adherence to the IRB-approved protocol and/or IRB policies and procedures. Monitoring plans should describe 1) who is responsible for file maintenance, 2) what will be maintained, 3) how often files will be reviewed and using what method, and 4) where documentation will be retained (for both Regulatory Documents and Participant files).

The original participant files including consent forms, enrollment log, adverse event form, and relevant Notes to File will be retained at the study sites with restricted access. The regulatory documents including sites protocol, IRB approval, staff training log and delegation log will be maintained at local study sites. Site PIs and study coordinators are responsible for file maintenance and reporting updates to the Broad project manager.

Once recruitment starts, the Broad project manager will check in with sites on a bi-weekly basis about the study progress, this includes any issues arising during screening, consenting process, enrollment and any reportable events. The electronic study enrollment log and any adverse events form will be sent to the Broad project manager and will be reviewed on a bi-weekly basis. Copy of the regulatory documents will be sent to the Broad when there's updates in IRB protocol status or staff changes. Sites IRB approvals will be reviewed and submitted to ESTR upon receipt. In addition to bi-weekly conference calls, the Broad project manager will go onsite visits to each study site per year to monitor compliance in consents process and record keeping. Both regulatory and participants' files will be reviewed during the site visits. If there's any findings, the project manager will re-train site researchers, and report the results to HSPH IRB.

15. Data or Biospecimen Sharing ☐ N/A. *Skip to next section.*

If you plan to establish a repository, please submit a separate application using the HLC Repository Protocol Template.

15.1. Describe the plan to send data/specimens to research collaborators outside of Harvard. ☐ N/A

We will follow the NIH Genomic Data Sharing Policy and NIMH policies on sharing data. All subjects will be explicitly consented for future research use and broad data sharing. The phenotypic data collection will be harmonized with other major psychiatric genomics projects in ancestrally diverse populations to the extent possible, e.g., the Populations Underrepresented in Mental Illness Association Studies (PUMAS), to maximize its utility in future studies. We will deposit the anonymized genotype and phenotypic data to NIMH funded National Data Archive (NDA), once the quality control procedures are completed, which we expect to take place within three months after the final data production in early Year 5 of the project. The data will be released within six months after the data submission or at the acceptance of initial publication, whichever occurs first.

We will publicly release genomic summary results (GSR) generated from our analyses, including the allele frequencies, effect size estimates, p-values and etc, both through an interactive web portal and as downloadable files hosted at the NHGRI-EBI GWAS Catalog. We will release GSR as soon as possible but no later than the acceptance of initial publication.

We will also contribute all the data to the Psychiatric Genomics Consortium (PGC) Bipolar Workgroup and the Bipolar Sequencing Project (BSC) to further increase the power for psychiatric genomics researches. We will make the contribution no later than the acceptance of the initial publication.

We will identify where the data will be available and how to access the data in any publications and presentations that we author or co-author using these data. As we will be using the NDA, which is an NIH- funded repository, this repository has policies and procedures in place that will provide data access to qualified researchers, fully consistent with NIH data sharing policies and applicable laws and regulations. These data will be shared with investigators working under an institution with a Federal Wide Assurance (FWA) and could be used for secondary study purposes such as finding genes that contribute to the risk architecture of bipolar disorder. The names and Institutions of persons either given or denied access to the data, and the bases for such decisions, will be summarized in the annual progress report.

15.2. Describe the plan to receive data/specimens <u>from</u> collaborators outside of Harvard.
☐ **N/A**

Biological samples from Taiwan, Singapore and Japan will be banked at the Broad Institute Genomics Platform; biological samples from South Korea and India will be banked at their respective facilities due to local regulations that either prevent them from leaving their country or origin, or prevent them from being banked in a foreign country. We will maintain a centralized sample manifest so that all samples can be located, tracked and be requested for future studies with proper legal and compliance approvals.

Phenotyping data will be collected and saved in REDCap database. The main electronic database for clinical and demographic information will be used to centrally manage, store, and transfer our phenotypic and genotypic data. At each site, the electronic database on which the data is recorded will be compatible with the main database. The database at each site will be password protected with access restricted to authorized members of the research team. Each site will use the same data entry software, which will be provided by Broad Institute at the initiation of the study. Training and education in the use of the software and database will be conducted by Broad staff during the initial site training visits and through monthly phone contact meetings. The data collected at each site will be securely transmitted on a weekly basis to the main database. Data at the main site will be reviewed weekly and available for discussion at the monthly phone meetings. No personally identifying information will be

entered into the central database. All participants will be assigned a subject identification number. All data will be identified only by this number.

16. Clinical Trials ☒ **N/A.** *Skip to next section.*
Complete this section for clinical trials, including NIH funded clinical trials or applicable clinical trials (ACT) under the FDA Amendments Act. To determine if a study meets the definition of a clinical trial, follow the guidance in the "Preparing the Research Protocol" section of the Investigator Manual.

16.1. Describe how this study meets the definition of a clinical trial.

16.2. Describe plans for registering this project in a clinical trials registry, e.g., clinicaltrials.gov. If available, provide the registry record number.

16.3. Describe plans for posting the clinical trial consent form on a publicly available federal website per federal requirements in the Common Rule (45§46.116(h)).

17. Device *This section should be completed if the study involves the use of any device on/in/with human subjects, and/or the use any device utilizing human specimens, which meets the FDA definition of a medical device.* ☒ **N/A.** *Skip to next section.*

17.1. Describe the device, including the generic or common name, brand name (if applicable), purpose, function/operation, and whether it is an implant. Indicate who is providing this device for research use.

17.2. Indicate the FDA status of the device as it is being used for the proposed research:
☐ FDA-approved device being used "on-label" (i.e., FDA-approved purpose, population, manner).
☐ FDA-approved device that is being used "off-label" (i.e., for a different purpose, population, or in a different manner than approved).
☐ Not approved by the FDA.

17.3. Indicate the IDE Status of this device:
☐ The use of this device has an IDE.
☐ The use of the device qualifies for an Abbreviated IDE.
☐ The use of the device is exempt from the IDE requirements.

17.4. Has the FDA made a determination as to whether the device is Significant Risk or Non-Significant Risk? ☐ No ☐ Yes: *If yes, indicate the FDA's determination.*

17.5. Describe plans for storage control, and dispending of the product so that (1) only authorized investigators will use the product; (2) the product will only be used in participants who have provided consent, and (3) there will be documented tracking of each product, including unique identifiers and any return/disposal.

18. Drug/Biologic *This section should be completed if the study involves the use of any drug/biologic on/in/with human subjects which meets the FDA definition of a drug/biologic.* ☒ **N/A.** *Skip this section.*

18.1. Describe the drug or biologic, including the generic or common name, brand name (if applicable), dosing, route of administration, number of doses, timing of administration. Indicate who is providing the drug, biologic, supplement for research use.

18.2. Indicate the IND Status of this drug or biologic and who holds the IND: (select one)

☐	There is an IND approval from the FDA for the use of this item.
	The IND is held by:
	The IND number is:
☐	An IND application has been, or will be, submitted to the FDA.
	The IND will be held by:
	The IND number is:
☐	An IND approval is not required.

18.3. Describe how dispensing, delivery and administration will be performed, and by whom. Include information about control (e.g., locked storage), tracking (e.g., lot number, returned pills), documentation, storage, and return/disposal.