There has recently been marked progress in identifying genetic risk factors for major depression (MD) and bipolar disorder (BD); however, few systematic efforts have been made to elucidate heterogeneity that exists within and across these diagnostic taxa. The Affective disorders, Environment, and Cognitive Trait (AFFECT) study presents an opportunity to identify and associate the structure of cognition and symptom-level domains across the mood disorder spectrum in a prospective study from a diverse US population.

Participants were recruited from the 23andMe, Inc research participant database and through social media; self-reported diagnosis of MD or BD by a medical professional and medication status data were used to enrich for mood-disorder cases. Remote assessments were used to acquire an extensive range of phenotypes, including mood state, transdiagnostic symptom severity, task-based measures of cognition, environmental exposures, personality traits. In this paper we describe the study design, and the demographic and clinical characteristics of the cohort. In addition we report genetic ancestry, SNP heritability, and genetic correlations with other large cohorts of mood disorders.

A total of 48,467 participants were enrolled: 14,768 with MD, 9864 with BD, and 23,835 controls. Upon enrollment, 47% of participants with MD and 27% with BD indicated being in an active mood episode. Cases reported early ages of onset (mean = 13.2 and 14.3 years for MD and BD, respectively), and high levels of recurrence (78.6% and 84.9% with >5 episodes), psychotherapy, and psychotropic medication use. SNP heritability on the liability scale for the ascertained MD participants (0.19–0.21) was consistent with the high level of disease severity in this cohort, while BD heritability estimates (0.16–0.22) were comparable to reports in other large scale genomic studies of mood disorders. Genetic correlations between the AFFECT cohort and other large-scale cohorts were high for MD but not for BD. By incorporating transdiagnostic symptom assessments, repeated measures, and genomic data, the AFFECT study represents a unique resource for dissecting the structure of mood disorders across multiple levels of analysis. In addition, the fully remote nature of the study provides valuable insights for future virtual and decentralized clinical trials within mood disorders.

Translational Psychiatry (2022)12:121; https://doi.org/10.1038/s41398-022-01877-2

INTRODUCTION

Mood disorders have a high lifetime prevalence in the general population and represent the leading cause of disability worldwide [1, 2]. Moreover, mood disorders cause marked impairiment in social and occupational functioning, resulting in a high burden for the individual and to society [3, 4]. Twin and family studies show moderate-high heritability for these syndromes, indicating a prominent role for genetic variation in conferring susceptibility [5–8]. MD has a lifetime prevalence of 15% [9] and twin-heritability of 30–40% [5, 10]. In contrast, BD has a lifetime prevalence of 2.4% [11] and twin-heritability ~70% [6, 12]. Genomic analyses have shown that mood disorders are highly polygenic with likely thousands of small-effect loci contributing to susceptibility [13, 14]. Significant progress has been made in identifying common genetic risk variants associated with MD and BD, most recently from the Psychiatric Genomics Consortium (PGC). The PGC Bipolar working group identified 40 independent BD loci in a sample of 40,000 BD cases [15], and the PGC MD working group identified 102 independent loci associated with MD from more than 246,000 cases [16]. Despite these successes, a major obstacle in psychiatric genetics is our inability to map these signals to the symptom patterns, cognitive deficits and maladaptive decision-making that characterize mood disorders.

Maria Dalby 1, 2, 5–7, Morana Vitezic 1, Niels Plath 1, Lene Hammer-Helmich 1, Yunxuan Jiang 3, Chao Tian 3, Devika Dhamija 3, Catherine H. Wilson 3, David Hinds 3, 5, 7, 8, 9, 23andMe Research Team*, Patrick F. Sullivan 2, 4, Joshua W. Buckholtz 5, 6, 9 and Jordan W. Smoller 7, 8, 9

© The Author(s) 2022, corrected publication 2022
One critical open question is how genetic risk affects human cognition to predispose the development of mood disorder symptoms and related behaviors. With up to 90% of patients with major depression (MD) or bipolar disorder (BD) exhibiting impairment in multiple domains of cognition, this represents an important diagnostic and symptomatic feature in mood disorders and a key determinant of functional recovery [17, 18]. Much of the morbidity and mortality in mood disorders is due to behavioral factors, such as substance abuse, aggression, self-harm, and risky sexual behavior [19–21]. These behaviors, in turn, are thought to result from deficits in cognitive processes related to cost-benefit decision-making, reinforcement learning, social cognition, and executive function [22]. Many groups have reported phenotypic associations between mood disorders and some of these cognitive processes [23, 24]. However, such studies are typically small in size, limited in scope, and genetically uninformative, limiting insight into the underlying causes of cognitive dysfunction and maladaptive behavior in mood disorders.

It is widely recognized that the DSM-based nosology of psychiatric illness poorly captures two important features of mental disorders: the high degree of comorbidity between diagnostic taxa, and the profound symptom-level heterogeneity that exists within a given diagnostic taxon [22, 25–27]. These features suggest the existence of latent transdiagnostic symptom clusters in mood disorders and are consistent with evidence for shared genetic liability between otherwise categorically distinct psychiatric disorders [28–33]. To date, we know little about how much of the shared variance among mood disorder symptoms, cognitive function and maladaptive behavior is due to genetic factors. Likewise, GWAS estimate the proportion of variance in liability attributable to common variants genome-wide (SNP-heritability) to be ~9% for MD and 18% for BD [15], which are fractions of the pedigree-based estimated heritability. This accords with the significant role of non-genetic factors in mood disorder risk. In particular, a number of environmental risk factors have been identified for mood disorders, including poverty and traumatic life events, particularly in early life. Understanding the mechanisms through which such environmental influences interact with genetic susceptibility is key to elucidating the risk architecture of mood disorders. However, existing GWAS data sets are unable to answer these and other important open questions because of practical constraints that preclude the collection of an appropriately rich set of phenotypic data at scale.

To bridge these gaps, we leveraged technological advances in web-based participant recruitment, diagnostic assessment and cognitive testing to create the AFFECT study. The AFFECT study employed a longitudinal case-control design in nearly 50,000 US-based participants with BD, MD, and controls. Study participants were recruited from the 23andMe, Inc research participant database and through social media, representing a diverse sample given at one or two time points as noted in Table 1.

For each GWAS, we restrict participants to a set of individuals who had a phenotype status determined through an analysis of local ancestry estimation [34] and a maximal set of unrelated individuals was chosen for each GWAS analysis using a segmental identity-by-descent (IBD) estimation algorithm [35].

**Methods**

**Cohort design**

This genetic, case-control study was designed to enroll three cohorts: 15,000 participants with MD, 10,000 participants with BD, and 25,000 controls with no lifetime MD or BD. Of these, 1533 participants (3.06%) withdrew consent or failed to return the spit kit or intake survey before the study termination date and were excluded. Participant eligibility criteria were: age between 18 and 50 years upon enrollment; residence in the United States; access to a desktop or laptop computer; and no reported diagnosis of Parkinson’s disease, essential tremor, schizophrenia, or Alzheimer’s disease. Enrollment required that the participants self-reported having been diagnosed with MD or BD by a medical professional and prescribed medication to treat such a disorder. Enrollment into the control cohort required that participants reported no lifetime diagnosis of BD, MD, generalized anxiety disorder, or post-traumatic stress disorder (PTSD) as well as never having been prescribed an antidepressant, mood stabilizer, or antipsychotic medication. All study participants had to provide informed consent and a saliva sample for SNP array genotyping, and be willing to complete the online study sessions over the course of 9 months. The study was conducted between August 2017 and September 2019 and online recruitment of participants, genotyping, and survey data collection were performed by 23andMe. Figure 1 illustrates the enrollment flow and study setup. Participants were recruited through two channels: all controls and approximately one-fifth (n = 4997) of all case participants were recruited from 23andMe’s existing customer database through email or logged-in website invitation. All other case participants (n = 9635) were recruited through social media such as Facebook and enrolled as new 23andMe customers. Study participants who met the eligibility criteria received compensation depending on if they were existing or new 23andMe customers. Existing customers, who had purchased a 23andMe kit prior to joining the study, received a $20 Amazon gift card. New customers received the 23andMe Health + Ancestry Service, including a DNA test kit, at no cost.

**Study assessments**

The study content was designed by the AFFECT investigators and administered by 23andMe. The self-reported survey and test battery (Table 1) was initiated at session 1 with an extensive background survey covering: demographics (i.e., age, gender, race, ethnicity), socioeconomic information (i.e., marital status, current employment, education, parental education, income), clinical details about the given disorder (cases only; e.g., age of onset, current and past episode characterization), family psychiatric history, the Self-rated Diagnostic and Statistical Manual of Mental Disorders (DSM-5) Level 1 Cross-Cutting Symptom Measure, and adverse childhood experiences (scale and scoring details in Supplementary Materials).

The mood and medication survey was also given at session 1 and repeated in sessions 2–5, 7, and 9. This survey included: medication history (session 1), changes in medications (all follow-up surveys), life events/lifestyle (e.g., alcohol use and sleep patterns), Altman Self-rating of Mania (ASRM) scale, and Patient-Reported Outcomes Measurement Information System (PRO-MIS)-Depression scale (scale and scoring details in supplementary materials). The study battery further included standardized behavioral tasks assessing risk, impulsivity and psychopathic traits and five cognitive tools designed to assess different domains of functioning. The cognitive tests were either given at one or two time points as noted in Table 1.

**SNP genotyping**

We evaluated common variant genetic contributions to risk for MD and BD using SNP array data. DNA extraction and genotyping were performed on saliva samples by the National Genetics Institute, a CLIA-licensed clinical laboratory and a subsidiary of the Laboratory Corporation of America. Samples were genotyped, phased and imputed by 23andMe standardized pipeline, as described in detail in Supplementary Materials. Roughly 9.22 million high-quality genotyped and imputed SNPs on autosomal and X chromosomes were tested.

For each GWAS, we restrict participants to a set of individuals who had a specified ancestry determined through an analysis of local ancestry estimation [34] and a maximal set of unrelated individuals was chosen for each GWAS analysis using a segmental identity-by-descent (IBD) estimation algorithm [35].

**Genome-wide associations**

GWAS was performed on MD versus controls, BD versus controls, mood disorder (MD, BD) versus controls and MD versus BD using a logistic
regression model: \( \text{case/control } \sim \text{ age } + \text{ sex } + \text{ top 5 Principal Components (PCs)} + \text{ genotyping platforms } + \text{ genotype} \). GWAS was first performed separately on individuals of European, African American, East Asian, Latino ancestry, and combined by fixed-effect meta-analysis using METAL [36]. GWAS results were adjusted for the genomic control inflation factors, which can be found under each Manhattan plot in Supplementary Figures. Note that the study enrollment channel (existing/enrolled customers) was embedded in the genotype platform term, where around 80% of existing customers were genotyped on 23andMe’s genotype platform v4, while all newly enrolled participants were genotyped on platform v5 (Supplementary Table 1). Across all results, we removed SNPs that had an available sample size of less than 20% of the total GWAS sample size; where logistic regression results that did not converge due to complete separation, identified by absolute value of effect size or standard error greater than 10 on the log-odds scale; or that had MAF < 0.1%.

**SNP-heritability and genetic correlations**

We used LD score regression (LDSC) [37] v1.0.1 to estimate SNP-heritability \( (h^2_{SNP}) \) from GWAS summary statistics for European ancestry MD and BD including variants with \( r^2 > 0.8 \) and minor allele frequency \( \geq 0.01 \). Estimates of \( h^2_{SNP} \) on the liability scale depend on the assumed lifetime prevalence of each disorder in the population \( (K) \). We report \( h^2_{SNP} \) with \( K = [0.001 – 0.3] \) for MD and \( K = [0.001 – 0.03] \) for BD.

Genetic correlations \( (r_g) \) to external summary statistics were also performed using LDSC [37]. External data included: the PGC MDD meta-analysis samples PGC-MDD1 (2013) [10], PGC-MDD2 excluding the 23andMe sample (2018) [38], and PGC-MDD3 excluding the 23andMe sample (2019) [16]; the 23andMe discovery sample of MDD (herein Hyde et al., 2016; where 5.0% of MD cases and 4.3% of controls from the AFFECT study were also included in Hyde et al.) [39]; the two most recent PGC BD meta-analysis samples PGC-BD2 (2019) [40] and PGC-BD3 (including the PGC-BD3 type I and type II sub-cohorts (2020) [15]); the most recent PGC SCZ meta-analysis samples PGC-SCZ2 (2014) [41] and PGC-SCZ3 (2020) [41, 42]. Data was obtained from https://www.med.unc.edu/pgc/download-results/ and through the 23andMe data-access portal.

**Statistical analyses**

Sample comparisons were conducted using R (v3.5.2). Descriptive statistics were performed on the total participation pool and on subgroups: the three cohorts of MD, BD, and controls; within subtype of BD diagnosis (BD1 vs. BD2) and, within each cohort subgroups based on enrollment strategy (i.e., participants drawn from the 23andMe database and participants enrolled through social media for this study). For categorical variables, the number and percentage were reported for each value. For quantitative variables, the mean, median, standard deviation, and ranges were reported. Differences in demographic and clinical covariates were compared using regression models (continuous or categorical variables) and Fisher’s exact tests for categorical variables.

**RESULTS**

Cohort characteristics

A total of 48,467 participants were included in these analyses: 14,768 reported that they had been diagnosed and treated for MD, 9864 had been diagnosed and treated for BD, and 23,835 were controls with no lifetime history of MD or BD (Fig. 1). The BD cohort contained 3070 (31.2%) BD subtype I (BD1), 5053 (51.3%) BD subtype II (BD2), and 1718 (17.5%) did not specify the latest type of BD diagnosis received (BD unspecified-type). Among all...
participants, 72% were female and the mean age was 32.3 years (range 18–52 years). Most participants were of European ancestry (71.9%) followed by Latino (14.2%), African American (3.8%), and East Asian (3.6%) ancestry (Table 2).

Participant completion rates ranged from 28 to 100% (mean 42.6%) per session and were lower for cognitive assessments than for surveys (Supplementary Table 2). Study retention (i.e., number of assessments completed) was highest for MD cases (mean 50.2%, SD 30.6) followed by BD cases (mean 45.2%, SD 30.6), lowest for controls (mean 38.2%, SD 28.8), and higher in females (mean 45.0%, SD 30.1) than males (mean 38.7%, SD 29.9) (Supplementary Fig. 1). Study retention was positively correlated with educational level and age, and negatively correlated with reported adverse childhood experience score, BMI, ASRM score,
and the DSM-5 cross-cutting domains of substance use, anxiety, depression, anger, suicidal ideation, and sleep problems (Supplementary Fig. 2).

Marital status, highest education achieved, and current socioeconomic status were reported at baseline and followed-up by a brief status assessment during each longitudinal assessment. Overall, socioeconomic status was significantly lower for cases, especially BD participants (Supplementary Table 3). In particular, we found that 19.5% and 26.9% of MD and BD participants, respectively, were currently not in paid employment as compared to only 7.3% of the control cohort. We observed an ascertainment effect in which case participants drawn from the 23andMe database (existing consumers) showed higher yearly salary and educational level than those enrolled through social media (multivariate analysis, $P < 1.0 \times 10^{-16}$). After adjusting for enrollment method, however, significant socioeconomic differences remained between cases and controls (multivariate analysis, $P < 1.0 \times 10^{-16}$, Supplementary Table 3).

**Disease history**

Figure 2A and Supplementary Table 4 summarizes the clinical features of MD and BD cases and highlights that both MD and BD presented with high disease severity. Most participants reported symptom onset in adolescence (MD; mean 13.2 (SD 5.1), BD; mean 14.3 (SD 5.2)) while formal psychiatric diagnosis was not typically received until early adulthood (MD; mean 19.5 (SD 6.6), BD; mean 23.2 (SD 7.6)), consistent with prior studies [43–45]. The course of illness differed between the disorders; BD cases tended to report short but recurrent episodes: 52.2% of the participants had experienced $>10$ episodes and 80.0% reported a typical episode duration of $<3$ months. In contrast, MD cases had fewer episodes of longer duration: 59.7% had experienced $\geq 10$ episodes, 47.7% reported a typical episode duration of 3–6 months or longer, and 10.0% reported episode duration $\geq 1$ year (Fig. 2A, Supplementary Fig. 3A, B).

As expected, psychotropic medication use was common, since this was an inclusion criterion: 23,202 (96.4%) of cases reported having taken medication for a mood disorder in the prior 5 years, 17,292 (70.2 %) were taking medication at baseline, and 7726 (31.4%) began or restarted a medication during the study. MD and BD participants (respectively) reported use of the following treatments in the prior 5 years and/or at present: antidepressants 13,803 (95.4%) and 8508 (88.1%); mood stabilizers 4875 (33.8%) and 8394 (86.9%); antipsychotics 6133 (42.4%) and 7972 (82.5%); and electroconvulsive therapy 107 (0.9%) and 144 (2.0%). Most cases had received cognitive or behavioral psychotherapy in the past 5 years (MD 9770, 67.7%, BD 7235, 79.3%; Supplementary Fig. 3C, D), most commonly 1–2 times a week. BD1 cases had the highest rates of symptom-related hospitalization (63.6%), although the rates were also high for the other mood disorder diagnoses (BD2, 46.1%; MD, 29.0%).

**Symptom state**

Nearly half of the MD cases ($N = 6971, 47.5%$) and about a quarter of the BD cases (2729, 27.8%) reported that they were experiencing an episode at baseline (Table 3, Supplementary Fig. 4). Most BD participants reported their current episode as depressive (1694, 62.7%). A current manic episode was reported in 219 (7.13%) BD1 participants, 106 (6.17%) unspecified-type BD participants, and a current hypomanic episode was reported across BD type: BD1 140 (16.0%), BD2 407 (29.4%), and 55 (11.9%) unspecified-type. We further observed that participants enrolled through social media exhibit greater disease burden (Supplementary Fig. 5, Supplementary Table 4) and were more likely to be in active mood episode compared to participants drawn from the 23andMe research participant database (41.0% versus 34.0%).

We defined probable depressive episodes using the Level 1 DSM-5 cross-cutting measure—depressive domain ($T$-score $\geq 60$), which identified 71.7% of all cases being in a depressive episode at baseline. Additionally, we defined a probable manic/hypomanic episode from the Level 1 DSM-5 cross-cutting measure—manic domain ($T$-score $\geq 2$) and the ASRM scale ($T$-score $> 5$), identifying 28.5% of BD participants being in an episode at baseline (Table 3). When comparing the symptom scale-based episodes with the self-identified episodes at baseline, we found reasonable correspondence for depressive episodes ($\kappa = 0.43$ and $\kappa = 0.36$ respectively for MD and BD) and a more modest correspondence for manic or hypomanic episodes ($\kappa = 0.22$).

**Symptom-level comorbidities**

The DSM-5 self-rated cross-cutting symptom measure assesses 13 transdiagnostic symptom domains of relevance across psychiatric...
Table 3. Self-identified episode (i.e. “Are you currently experiencing an episode?”), “What type of episode are you experiencing?” and symptom scale-based episode of cases at baseline.

|                          | Self-identified episode (A), n (%) | Symptom-scale episode (B), n (%) | Self-identified and Symptom-scale episode (A∩B), n (% A | B, % B|A) | Cohen’s κ |
|--------------------------|-----------------------------------|----------------------------------|--------------------------------------------------------|----------------------------------------|----------|
| **MD**                   |                                   |                                  |                                                        |                                        |          |
| Depressive               | 6971 (47.5)                      | 8575 (59.3)                      | 5980 (87.5, 70.1)                                      | 0.43 (±0.01)                           |          |
| **BD**                   |                                   |                                  |                                                        |                                        |          |
| Depressive               | 1694 (17.3)                      | 5502 (57.0)                      | 1637 (93.5, 71.4)                                      | 0.36 (±0.04)                           |          |
| Manic/hypomanic          | 927 (9.5)                        | 2754 (28.5)                      | 613 (66.7, 22.4)                                      | 0.22 (±0.02)                           |          |

The intersection (A∩B) shows number and proportion overlap, where A|B: self-report given symptom-based outcome, B|A: symptom-based outcome given self-report. Cohen’s κ given with 0.05 confidence interval (CI). N based on * = 14,447,  = 14,690,  = 9663,  = 9802.

diagnosis [34] (scoring details given in Supplementary Materials). We found that both MD and BD participants exhibited a wide-range of transdiagnostic symptoms (median number of positively screened symptom domains = 9), a clear distinction to the control cohort (median number of positively screened symptom domains = 2) (Fig. 2B, Supplementary Table 5). The most common symptom domains in cases were depression, mania, somatic symptoms (i.e. aches and pains), and anxiety. Furthermore, sleep problems and substance use symptoms provided the strongest differentiation of BD from MD (multivariable analysis, coefficient 0.43 (95% CI ±0.03) \( P < 2.2e^{-16} \), coefficient 0.41 (95% CI ±0.05) \( P < 2.2e^{-16} \), respectively).

Regarding non-psychiatric conditions, MD and BD participants reported higher rates of comorbidities compared to controls. This was particularly evident for inflammatory and neurological disorders (multivariable analysis, OR ≥ 3.03 \( P < 0.001 \), Supplementary Table 3).

**Family psychiatric history**

Family history prevalence of anxiety disorder, MD, BD, or PTSD in first-degree relatives is shown in Supplementary Table 6. Rates were significantly higher for all disorders among cases (78.4 %) compared to controls (Fisher’s exact OR = 4.2 (95% CI ±0.1), \( P < 2.2e^{-16} \)), particularly for the same disorder and within BD subtypes (Fisher’s exact OR (95% CI) MD − 6.6 (0.6), BD1 − 3.1 (±0.4), OR − 5.0 (±1.0), \( P < 2.2e^{-16} \), Supplementary Fig. 6). The prevalence of mental disorders in first-degree relatives of controls (33.0%) was comparable to rates reported in population-based samples [46].

**Environmental influences**

Reported adverse childhood experiences (ACE) were assessed across multiple domains (i.e., psychological and sexual abuse, neglect, and household dysfunction) [47, 48]. Childhood adversity was common, with 63.9% of participants reporting at least one ACE. The total ACE score was significantly higher in cases than controls, with almost twice as many ACEs reported (case mean = 3.96, control mean = 2.00, \( P < 1.0 \times 10^{-16} \)). Moreover, BD cases reported more ACEs than MD cases (Supplementary Table 7). Within ACE domains, physical and emotional neglect showed the largest association with mood disorders (OR = 5.6, 95% CI ±0.4); again, these associations were considerably stronger in BD cases (OR = 6.54, 95% CI ±0.34).

**SNP-heritability and genetic comparability**

GWAS was conducted in European ancestry participants for mood disorder (MD + BD), each disorder separately, BD subtypes, and comparing MD versus BD. Furthermore, a trans-ethnic meta-analysis of European, Latino, African American and East Asian GWAS was conducted for MD and for BD. Variant-level analysis, which was not the focus of this paper, is provided in Supplementary Figs. 7–22 and sample sizes for each GWAS can be found in Supplementary Table 8.

**DISCUSSION**

The AFFECT study was initiated to advance our understanding of phenotypic and genetic heterogeneity in MD and BD and to clarify the role of shared genomic and environmental risk factors that may transcend their diagnostic boundaries. Several aspects of AFFECT are notable including the administration of task-based measures indexing multiple domains of cognition (e.g. executive, motivational, and social) that capture key facets of the Research Domain Criteria (RDoC) [50] framework; transdiagnostic symptom assays; the assessment of trait and environmental risk and resilience factors; and the repeated measures design enabling analysis of change in symptoms and multi-domain cognitive task performance. Here, we have presented baseline characterization of the cohort and summarized the clinical features of MD and BD cases.

The US-based study participants were ascertained from the general 23andMe participant database and from social media. Control participants did not self-report diagnosis of or treatment for mood disorders. Case participants self-reported a clinician-ascertained diagnosis of MDD or BD (I or II) and were currently using one or more prescribed medications to manage their symptoms. Additional study ascertainment criteria pertained to
age (18–50 years old) and the absence of Parkinson's disease, Alzheimer's disease, essential tremor, or schizophrenia diagnosis. Demographic and socio-economic features of BD and MD cases in the AFFECT study were largely comparable to those reported in epidemiologic and clinical samples [51–53] with a substantial female predominance among cases. Consistent with prior research [54, 55], reported adverse childhood experiences were relatively common and associated with significantly increased risk of mood disorder.

Prior studies have shown that selective participation represents a potential source of bias in both epidemiological and genetic association studies [56, 57]. Consistent with this, several features of the cohort differ from those seen in many clinically ascertained mood disorder cohorts. For example, educational attainment and income levels among MD cases were somewhat higher than reported in population-based samples [52] as might be expected given the ascertainment through a direct-to-consumer genomics company. Interestingly, we observed some differences within the sample: lower socioeconomic status and greater illness severity were observed among those recruited through social media compared to participants drawn from the existing 23andMe consumer database. Although it might be expected that cases recruited through direct-to-consumer genomics and social media platforms would have less burden of illness compared with those ascertained clinically, this was not the case. In fact, most mood disorder cases in this study reported early-onset illness, recurrent episodes, positive family history, and treatment with medication and psychotherapy. Indeed, a history of psychiatric hospitalization among MD cases was higher (29%) than that reported in a representative sample of US adults (12%) [52]. Together, these suggest a high disease burden (significant impairment and dysfunction) in our cohort.

Overall, 71.7% of AFFECT participants reported symptoms of a current depressive episode at baseline, and 28.1% of BD cases reported current manic or hypomanic symptoms. This likely reflects the fact that BD2 was overrepresented in our BD cohort (51.3%) relative to population-based samples [11, 53], but may also suggest that remote study participation is more likely for euthymic and depressive BD patients. We found that the agreement between self-reported and mood scale ratings for mania was limited. This underlines the limitations of self-reported assessments and symptom-based outcomes as discussed elsewhere [58].
Despite these considerations, we expect the AFFECT study to contribute importantly to understanding the genetic basis of mood disorders. The incorporation of transdiagnostic symptom and behavior measures, longitudinal symptom assessments, and task-based measures of neuro- and social cognition, make this a unique resource for genomic studies. In the initial GWAS of the AFFECT mood disorders, we identified several genome-wide significant loci; the strongest association was between MD and SNPs within NGR1, a gene encoding a synaptic adhesion protein that has been robustly associated with depression in prior studies [16, 59]. Recent analyses have found that GWAS of MD samples characterized by “minimal phenotyping” (e.g. based on self-report of prior diagnosis and/or treatment for depression) show lower heritability and are enriched for less specific genetic effects on MD compared with samples diagnosed using strict syndromal criteria [60]. In this context, it is notable that the estimated liability scale \( h^2_{\text{SNP}} \) for AFFECT MD (0.19–0.21) is in the same range as “strictly-defined lifetime MDD” in that analysis and higher than what is seen in broadly-defined MD cohorts, including the previous 23andMe self-reported depression cohort [16, 38, 39]. As demonstrated in previous work [61, 62], SNP heritability is a consequence of several known and unknown effects, including the exclusion of specific comorbidities, disease severity, and the use of controls from which other psychiatric disorders have been excluded [63].

Genetic correlation analyses indicate that AFFECT MD is highly correlated (\( r_5 = 0.71–0.85 \)) with MD ascertained in studies included in the PGC. Unexpectedly, however, genetic correlations between AFFECT-BD and published PGC GWAS of BD were relatively modest (\( r_5 = 0.38–0.43 \)) while the genetic correlation between the AFFECT MD and AFFECT BD was approximating 1. Indeed, the pattern of genetic correlations seen with AFFECT-BD closely resembled those of AFFECT-MD and did not vary between different BD samples, further underlining the importance of understanding heterogeneity within and across diagnostic taxa. The AFFECT study represents a unique cohort of remotely recruited individuals with MD and BD and controls. The availability of repeated measures over time as well as task-based cognitive domains will provide an important opportunity to examine the genomic basis of mood disorders and underlying traits. More in-depth analyses of these phenotypes and shared or unique contributions to BD and MD are forthcoming.

DATA AVAILABILITY
The top 10,000 SNPs for each GWAS are provided in Supplementary Tables 10–15. Participants provided informed consent and participated in the research online, under a protocol approved by the external AHRPP-accredited IRB, Ethical & Independent Review Services (E&I Review). Participants were included in the analysis on the basis of consent status as checked at the time data analyses were initiated. The full GWAS summary statistics for the 23andMe discovery data set will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit https://research.23andme.com/collaborate/#dataset-access/ for more information and to apply for access. Individual-level data are not publicly available due to participant confidentiality, and in accordance with the IRB-approved protocol under which the study was conducted. Researchers interested in the study’s individual-level data may apply to the 23andMe Research Innovation Collaborations program.

REFERENCES
1. Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N, et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet. 2011;43:977–85. https://doi.org/10.1038/ng.943
2. World Health Organization. Depression and Other Common Mental Disorders—Global Health Estimates. World Health Organization; 2017. p. 524.
3. Romera I, Perez V, Menchon JM, Craddock N, Hulley H, Polaiveja P, Gilaberte L. Social and occupational functioning impairment in patients in partial versus complete remission of a major depressive disorder episode. A six-month prospective epidemiological study. Eur Psychiatry. 2010;25:58–65. https://doi.org/10.1016/j.eurpsy.2009.02.007.
4. Rosa AR, Reinares M, Michalak EE, Bonnin CM, Sole B, Franco C, et al. Functional Impairment and Disability in Mood States in Bipolar Disorder. Value Heal.2010;13:984–98. https://doi.org/10.1111/j.1524-733X.2010.00768.x.
5. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: Review and meta-analysis. Am J Psychiatry. 2000;157:1552–62. https://doi.org/10.1176/appi.ajp.157.10.1552.
6. McGuffin P, Rijsdijk F, Andrew M, Sham P, Katz R, Cardno A. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. Arch Gen Psychiatry. 2003. https://doi.org/10.1001/archpsyc.60.5.497.
7. Kendler KS, Gatz M, Gardner CO, Pederson NL. A Swedish national twin study of lifetime major depression. Am J Psychiatry. 2006;163:109–14. https://doi.org/10.1176/appi.ajp.163.1.109.
8. Smoller JW, Finn CT. Family, Twin, and Adoption Studies of Bipolar Disorder. Am J Med Genet Semin Med Genet. 2003;123:C48–58. https://doi.org/10.1002/ajmg.c.20013.
9. Bremen E, Andrade LH, Hwang I, Sampson NA, Alonso J, de Girolamo G, et al. Cross-national epidemiology of DSM-IV major depressive episode. BMC Med. 2011. https://doi.org/10.1186/1741-7015-9-90.
10. Sullivan PF, Daly M, Ripke S, Lewis CM, Wray NR, Hamilton SP, et al. A meta-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry. 2013;18:497–511. https://doi.org/10.1038/mp.2012.21.
11. Merikangas KR, Jin R, He JP, Kesder RC, Lee S, Sampson NA, et al. Prevalence and correlates of bipolar spectrum disorders in the World Mental Health Survey Initiative. Arch Gen Psychiatry. 2011. https://doi.org/10.1001/archgenpsychiatry.2011.12.
12. Edvardsen T, Torgersen S, Røysamb E, Lygren S, Skre I, Onstad S, et al. Heritability of bipolar spectrum disorders. Unity or heterogeneity? J Affect Disord. 2008. https://doi.org/10.1016/j.jad.2007.07.001.
13. Sullivan PF, Daly MJ, Ripke S, Lewis CM, Lin DY, Wray NR, et al. A meta-Analyisis of genome-wide association studies for major depressive disorder. Mol Psychiatry. 2013;18:497–511. https://doi.org/10.1038/mp.2012.21.
14. Neale BM, Sklar P. Genetic analysis of schizophrenia and bipolar disorder reveals polygenicity but also suggests new directions for molecular interrogation. Curr Opin Neurobiol. 2015;30:131–8. https://doi.org/10.1016/j.conb.2014.12.001.
15. Mullins N, Forstner AJ, Os K, Sloofman LG, Steinberg S, Trubetskov Y, et al. Genome-wide association study of over 40,000 bipolar disorder cases provides novel biological insights. medRxiv. 2020;17:202. https://doi.org/10.1016/j.neurosci.2019.12.036-7.
16. Zuckerman H, Pan Z, Park C, Brietzke E, Musial N, Shariq AS, et al. Recognition and Treatment of Cognitive Dysfunction in Major Depressive Disorder. Front Psychiatry. 2019;1:11. https://doi.org/10.3389/fpsyg.2018.00655.
17. Zubrieta JK, Huguelet P, O’Neill RL, Giordani BJ. Cognitive function in euthymic Bipolar I Disorder. Psychiatry Res. 2001;102:9–20. https://doi.org/10.1016/s0165-1781(01)00242-6.
18. Fagiolini A, Fornione R, Macciari M, Cuomo A, Morana B.Dell’Osso B, et al. Prevalence, chronicity, burden and borders of bipolar disorder. J Affect Disord. 2013;148:161–9. https://doi.org/10.1016/j.jad.2013.02.001.
19. Hirschfeld RMA, Cass AR, Holt DCL, Carlson CA. Screening for bipolar disorder in patients treated for depression in a family medicine clinic. J Am Board Fam Pract. 2005;18:233–9. https://doi.org/10.3122/jabfm.18.4.233.
20. Kleinman LS, Lowin A, Flood E, Gandhi G, Edgell E, Revicki DA. Costs of bipolar disorder. Pharmacoeconomics. 2003;21:601. https://doi.org/10.2165/00019053-200321090-00001.
21. Goldberg JD, Chengappa KNR. Identifying and treating cognitive impairment in bipolar disorder. Bipolar Disord. 2009;11:123–37. https://doi.org/10.1111/j.1399-5618.2009.00716.x. SUPPL. 2.
22. Quraishi S, Frangou S. Neuropsychology of bipolar disorder: A review. J Affect Disord. 2002;72:209–26. https://doi.org/10.1016/s0165-0327(02)00091-5.
71. Peters J, Büchel C. The neural mechanisms of inter-temporal decision-making: Understanding variability. Trends Cogn Sci. 2011;15:227–39. https://doi.org/10.1016/j.tics.2011.03.002.
72. McClure SM, Laibson DI, Loewenstein G, Cohen JD. Separate neural systems value immediate and delayed monetary rewards. Science. 2004;306:503–7. https://doi.org/10.1126/science.110097.
73. Kable JW, Glimcher PW. The neural correlates of subjective value during inter-temporal choice. Nat Neurosci. 2007;10:1625–33. https://doi.org/10.1038/nn2007.
74. Rosenberg M, Noonan S, DeGutis J, Esterman M. Sustaining visual attention in the face of distraction: A novel gradual-onset continuous performance task. Attention, Perception, Psychophys. 2013;75:426–39. https://doi.org/10.3758/s13414-012-0413-x.
75. McIntyre RS, Best MW, Bowie CR, Carmona NE, Cha DS, Lee Y, et al. The THINC-Integrated Tool (THINC-it) Screening Assessment for Cognitive Dysfunction: Validation in Patients With Major Depression Disorder. J Clin Psychiatry. 2017;1–4. https://doi.org/10.4088/JCP.16m11329.
76. Lam RW, Saragoussi D, Danchenko N, Rive B, Lamy FX, Brevig T. Psychometric Validation of Perceived Deficits Questionnaire – Depression (PDQ-D) in Patients with Major Depression Disorder (MDD). Value Heal. 2013;16:A330. https://doi.org/10.1016/j.val.2013.08.046.
77. Lejuez CW, Richards JB, Read JP, Kahler CW, Ramsey SE, Stuart GL. et al. Evaluation of a behavioral measure of risk taking: The balloon analogue risk task (BART). J Exp Psychol Appl. 2002;8:75–84. https://doi.org/10.1037/1076-989X.8.2.75.
78. Baron-Cohen S, Jolliffe T, Mortimore C, Robertson M. Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. J Child Psychol Psychiatry Allied Discip. 1997;38:813–22. https://doi.org/10.1111/j.1469-7617.1997.tb01599.x.

ACKNOWLEDGEMENTS
We thank the project coordinator STF (23andMe Inc.), Lars-Peder Haahr (former employee at Lundbeck A/S), all AFFECT-study scientists from Lundbeck A/S, Massachusetts General Hospital, and 23andMe Inc. for valuable discussion and input. This work was supported by a Post Doc grant (8054-00026B) from the Innovation Fund Denmark (MD). Finally, we thank all study participants, who made this work possible.

AUTHOR CONTRIBUTIONS
Conceptualization: CHW, JWS, JWB. Data curation: MD, YJ, DD. Formal analysis: MD, YJ. Funding acquisition: 23andMe Research Team, NP. Project administration: 23andMe Research Team, NP. JWS, JWB. Supervision: PF5, LH-H, MV, DH, JWS, JWB. Writing the original draft: MD, PF5, JWS, JWB. Reviewing and editing: all authors.

23ANDME RESEARCH TEAM
Stella Aslibekyan3, Adam Auton3, Elizabeth Babalola3, Robert K. Bell3, Jessica Bielenberg3, Katarynna Bryc3, Emily Bulli33, Daniella Coker3, Gabriel Cuellar Partida3, Sayantan Das3, Sarah L. Elson3, Yonas Eshete3, James Foster3, Matthew G. Furlong3, Beatrice Gocza3, Kevin Golding3, Yeanling Han3, Adam Hiltunen3, Timothy Hoang3, Evan Ives3, Raghavendra Iyer3, Hiram Jaramillo3, Michael Jenkins3, Nathaniel Johnson3, Brian J. Jones3, Scott Kahan3, J. Kabcenell3, Animesh Khan3, Ray Knauf3, Kevin Kosco3, Tiffany Kwan3, Taylor LaMorte3, Joshua Lapi3, Julianne Laue3, Anna Lee3, Zainab Levison3, Eric Logsdon3, Sunny Lu3, Clayton Lueders3, Emily Luu3, Sarah L. Marotz3, Andrew Martinez3, Caleb Mathis3, Gary McCue3, William McGaw3, Yichen Mei3, Stuart McVean3, Brian McVean3, John Mennen3, Paul Mieczkowski3, Kali von Manteuffel3, Thayer Mann3, Cheri March-Leake3, Numerous members of the Genetics team3, Simon Moore3, Brandon Morey3, Adam Moseley3, David Mosher3, Adam Mountz3, Brie Neufer3, Michael Nihal3, Ryan Nomura3, Raja Nono3, Anil Nrichhi3, Shelley Noones3, Saurav Nogar3, Jack Opila3, Michael Pabst3, Steven Page3, Rachel Pan3, Joshua Parkinson3, Kamyar Parkani3, Charles R. Parikh3, Alejandro Paz Tovar3, J. Paul Pelletier3, Andrew Pearlman3, Garett Perdew3, Joseph Perkins3, Andrew Petkova3, Alyssa Pianalto3, Luka Plavetic3, Megan Pratt3, Murali Prasad3, Paul Préfontaine3, Erik Priest3, Jacob Price3, Kevin Prince3, Jordan Price3, Andrew Prins3, Eileen Pudlo3, Michael Puett3, Matthew Puskar3, Alexee Qiao3, Matthew Quinn3, Sarah Rafter3, Yoko Rajapakshe3, James Ray3, Jonathan Ready3, William Redman3, M. R.歼3, Jacob Ropicki3, Ralph Rose3, Parth Rowlingson3, Caroline Ruzicka3, Tobin Sacksteder3, Evelyn Salazar3, Joshua Salazar3, Ted Salomon3, Santhosh Samanta3, Ryan Sarmiento3, Alex Sater3, J. Michael Savell3, Nathaniel Schrader3, Anish Schmidt3, Ritesh Seshadri3, Wei Shih3, Colby Shnier3, Ashutosh Singh3, Ashley Sisti3, M. K. Singh3, Sarah Smith3, Michelle Smith3, Andrew Smith3, Richard Smith3, Matthew Smith3, Henry Sohn3, Danielle Spillman3, James Spindel3, Jake Spies3, Charles Spitaleri3, Adam Stahl3, William Stahl3, Christopher Stavros3, Andrew Steinfeld3, Alex Stortz3, Niko Szeto3, Andrew T., Steven T., Brian T., Andrew Tong3, Shiping Tong3, Ian Tong3, Jesse Tzeng3, Matthew Underdown3, Jon Underdown3, M. Underwood3, Michael Udell3, David Ullmann3, John Umland3, Neil Van Arnum3, Grant Varoquaux3, Aris Vanas3, Michael Varney3, Timothy Vasko3, Arman Verma3, Gaba Verma3, John Vickers3, Sarah Vitalis3, Darya V. Volkova3, Scott Voller3, Ian Waddell3, Shashank Wagner3, Lisa W. Wagger3, Linda Wagoner3, elevator3, Hunter Walsh3, Andrew Wander3, Paul Wang3, John Wang3, Andrew Wang3, Jessica Wang3, Evan Warfield3, Matthew Warren3, Timothy Warren3, David Wessels3, Andrew Westerfield3, Richard Whipple3, Alex White3, Michael Whisenand3, Tyler Wickett3, Wei Xiong3, Ryan Yoon3, Rakesh Yoon3, Will Young3, Andrew Yoshida3, James Young3, J. Yuen3, Allen Yu3, Mika Yust3, Lauren Zander3, John Zarnowski3, Emma Zawada3, Andrew Zheng3, Kristina Ziemba3, Jun Zhao3, Bo Zhang3, Peter Ziman3, Michael Zink3, Ryan Zobrist3, L. Zou3, Yujun Zou3, C. Zuo3.