Check for updates

ARTICLE

The addiction risk factor: A unitary genetic vulnerability characterizes substance use disorders and their associations with common correlates

Alexander S. Hatoum[1✉], Emma C. Johnson[1], Sarah M. C. Colbert[1], Renato Polimanti[2,3], Hang Zhou[2,3], Raymond K. Walters[4,5], Joel Gelernter[2,3,6,7], Howard J. Edenberg[8,9], Ryan Bogdan[10,11] and Arpana Agrawal[1,11]

Substance use disorders commonly co-occur with one another and with other psychiatric disorders. They share common features including high impulsivity, negative affect, and lower executive function. We tested whether a common genetic factor undergirds liability to problematic alcohol use (PAU), problematic tobacco use (PTU), cannabis use disorder (CUD), and opioid use disorder (OUD) by applying genomic structural equation modeling to genome-wide association study summary statistics for individuals of European ancestry (Total $N = 1,019,521$; substance-specific Ns range: 82,707–435,563) while adjusting for the genetics of substance use (Ns = 184,765−632,802). We also tested whether shared liability across SUDs is associated with behavioral constructs (risk-taking, executive function, neuroticism; Ns = 328,339−427,037) and non-substance use psychopathology (psychotic, compulsive, and early neurodevelopmental disorders). Shared genetic liability to PAU, PTU, CUD, and OUD was characterized by a unidimensional addiction risk factor (termed *The Addiction-Risk-Factor*, independent of substance use. OUD and CUD demonstrated the largest loadings, while problematic tobacco use showed the lowest loading. *The Addiction-Risk-Factor* was associated with risk-taking, neuroticism, executive function, and non-substance psychopathology, but retained specific variance before and after accounting for the genetics of substance use. Thus, a common genetic factor partly explains susceptibility for alcohol, tobacco, cannabis, and opioid use disorder. *The Addiction-Risk-Factor* has a unique genetic architecture that is not shared with normative substance use or non-substance psychopathology, suggesting that addiction is not the linear combination of substance use and psychopathology.

INTRODUCTION

Substance use and use disorders (SUDs) represent large and growing public health problems that account for nearly 6% of global disease burden [1]. SUDs, both licit and illicit, commonly co-occur with each other and also with non-substance psychopathology; comorbidity is associated with increased symptom severity [2] and worse outcomes (e.g., less responsivity to treatment, greater socioeconomic costs [3]). However, the etiology underlying shared risk across these disorders is poorly understood.

Shared genetic liability

According to twin studies, the moderate-large heritability (50–60%) of distinct SUDs (i.e., alcohol, nicotine, cannabis, and other illicit drugs) is partly attributable to a shared genetic vulnerability [4]. Similarly, genetic correlations estimated from genome-wide association study (GWAS) data support a shared

genetic vulnerability between SUDs (e.g., SNP-rG = .73 between alcohol use disorder and opioid use disorder) [5], between SUDs and substance use (e.g., SNP-rG = .78 between problematic alcohol use and drinks per week) [6], and between SUDs and psychopathology (e.g., SNP-rG = .33 between cannabis use disorder and major depressive disorder [7]). What remains unclear is the extent to which genetic liability across substance use *disorders* is *shared with* and *distinct* from that of substance *use* (but not dependence) and non-substance psychopathology, and what putative intermediate phenotypes may link shared genetic liability between SUDs and non-substance psychopathology.

Substance use and use disorder. Substance use and SUDs have substantial genetic overlap; however, genetic mechanisms that relate to SUD liability beyond normative or frequently occurring substance use remain. For opioids [8], alcohol [9–12],

[1]Washington University School of Medicine, Department of Psychiatry, Saint Louis, MO, USA. [2]Department of Psychiatry, Division of Human Genetics, Yale School of Medicine, New Haven, CT, USA. [3]Veterans Affairs Connecticut Healthcare System, West Haven, CT, USA. [4]Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [5]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [6]Department of Genetics, Yale School of Medicine, New Haven, CT, USA. [7]Department of Neuroscience, Yale School of Medicine, New Haven, CT, USA. [8]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA. [9]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA. [10]Department of Psychological & Brain Sciences, Washington University in St. Louis, St. Louis, MO, USA. [11]These authors contributed equally: Ryan Bogdan, Arpana Agrawal. ✉email: ashatoum@wustl.edu

and cannabis [7], the use and use disorder dimensions show differing associations with psychopathology (e.g., schizophrenia) and life outcomes (e.g., educational attainment) [7, 8, 10, 13].

Substance use and psychopathology. Recently, Lee and colleagues [14] identified three broad clusters (psychotic, compulsive, and early neurodevelopmental) representing shared and distinct genetic liability to eight non-substance psychiatric disorders. Further, polygenic liability to cross-diagnostic vulnerability is associated with substance use and SUDs [15]. In addition, ADHD shares genetic variance with substance use (e.g., cannabis use, nicotine use) and substance use disorders (e.g., problematic alcohol use) [16], and there is evidence that specific substance use disorder GWASs are correlated with a heterogeneous factor of autism, ADHD, and depression [17]. Collectively, these data suggest that substance use and use disorders share genetic liability with psychopathology.

Stage-based addiction individual differences and substance use disorders. SUD vulnerability has been conceptualized within a three-stage neurobiological model consisting of binge/intoxication, preoccupation/anticipation, and withdrawal/negative affect [18]. In this model, initial positive reinforcement is derived from stimulation of neural reward circuitry that drives impulsive behaviors in the context of under-developed tolerance. With continued use and progression towards SUD, the reinforcing properties of substances shift from positive to negative reinforcement; as use becomes compulsive, it functions to return the body to drug-present homeostasis and alleviate low mood, a predisposition to which is broadly indexed by neuroticism [19]. Following repeated drug-reward and drug-homeostasis pairings, cognitive preoccupation with the drug in expectation of reward/relief emerges in the context of impaired executive functioning [20]. While GWASs support genetic correlations between SUDs and risk-taking [5, 21], executive functioning [22], and negative affect [5, 21], the extent to which common genetic liability across SUDs relates to these constructs has yet to be examined.

The current study

Given evidence of shared liability to SUDs and other forms of psychopathology, understanding the shared and unique genetic contributions to SUDs and how these relate to heritable proxies for stage-based addiction constructs, non-substance psychopathology, and substance use may generate etiologic insights that improve psychiatric nosology, prevention, and treatment. To this end, we *first* estimate the shared genetic structure across SUDs by applying genomic structural equation modeling (gSEM) [23] to summary statistics generated by the largest GWAS of problematic alcohol use (PAU) [21], problematic tobacco use (PTU) [24, 25], cannabis use disorder (CUD) [7], and opioid use disorder (OUD) [5]. We name the shared variance across SUDs the *Addiction-Risk-Factor*. *Second*, we relate the *Addiction-Risk-Factor* to the genetics of behavioral constructs representing proxies of the stage-based model of SUDs. We estimate the extent to which genetic liability to risk-taking, executive function, and neuroticism are related to *The Addiction-Risk-Factor*. *Third*, we examine whether *The Addiction-Risk-Factor* is associated with the 3 factors representing genetic liability to non-substance psychopathology [14] (i.e., psychotic, compulsive, and neurodevelopment) and whether stage-based addiction constructs (i.e., risk-taking, executive function, neuroticism) indirectly link *The Addiction-Risk-Factor* to psychopathology. *Finally*, given that genetic liability to substance use (e.g., ever using, quantity-frequency) and later stages of SUDs are partially distinct [7, 10, 13], we repeat all analyses while incorporating genetic liability to substance use (i.e., alcohol drinks/week [25]; tobacco ever regularly use [25], cannabis ever use [26]) as covariates.

We hypothesized that SUDs and problem substance use would be largely characterized by a common genetic vulnerability (i.e.,

The Addiction-Risk-Factor) with evidence of potentially important substance-specific liability (e.g., metabolic and signaling pathways for a specific drug such as *ADH1B* variants with alcohol [27]). We hypothesized that (i) *The Addiction-Risk-Factor* would be associated with all three non-substance psychiatric clusters while retaining variance unique to itself, (ii) genetic liability to behavioral phenotypes representing vulnerability to stage-based addiction constructs (i.e., risk-taking, executive function, and neuroticism) would be associated with *The Addiction-Risk-Factor* and account for a proportion of the association between *The Addiction-Risk-Factor* and psychopathology factors, and (iii) after accounting for the genetics of substance use, *The Addiction-Risk-Factor* would retain unique variance (i.e., we expect significant residual genetic correlations among SUDs) and maintain similar patterns with non-substance psychopathology and stage-based constructs.

METHODS
Samples
Summary statistics from the largest available discovery GWASs were used to represent a genetic risk for each construct (more details are in Supplemental Table 1). These include: (i) 4 *SUDs* (problematic alcohol use [21], problematic tobacco use [24], cannabis use disorder [7], opioid use disorder [5]); (ii) 3 *substance use phenotypes* (alcohol drinks/week [25], lifetime ever smoking [25], lifetime cannabis use [26]); (iii) 3 *traits proxying the stage-based model of SUDs* (risk-taking, executive function, neuroticism); and (iv) 9 *non-substance psychiatric disorders*. Analyses were restricted to data from individuals of European ancestry because GWAS on these constructs in other ancestral origins are not available or are underpowered, and cross-ancestry analysis can confound genetic correlation estimates [28]. All GWAS summary statistics were filtered to retain variants with minor allele frequencies > 0.01 and INFO score > 0.90 for GSCAN and PGC [7, 25] and INFO score > 0.70 for the MVP [5, 29].

Problematic substance use/substance use disorder summary statistics
Problematic alcohol use. Summary statistics for problematic alcohol use (PAU) were derived from a meta-analysis of GWASs of DSM-IV alcohol dependence from the Psychiatric Genomics Consortium [11] (PGC-AD; $n = $ 11,569 case, 34,999 controls), ICD-9/10 based diagnoses of alcohol use disorders from the Million Veteran Program phase 1 and 2 data (MVP; $n = $ 45,995 cases; 221,396 controls) [9] and the Problem subscale score from the Alcohol Use Disorders Identification Test (AUDIT-P) [10] from the UK Biobank ($n = $ 121,604) [21]. The final GWAS summary statistics included data on 435,563 participants [21]. We also report on model fit with PGC-AD (instead of PAU) in the supplement (Alternative Models, M1).

Problematic tobacco use (PTU). We used summary statistics from the GWAS of the Fagerström Test for Nicotine Dependence [24] (FTND). As cigarettes per day is an item within the FTND and the genetic correlation between FTND and cigarettes per day is high (calculated rG = 0.97 CI = .12) [13], we combined CPD And FTND into a single indicator. We applied Multi-Trait Analysis of GWAS summary statistics (MTAG [30]) to summary statistics generated from the GWAS and Sequencing Consortium of Alcohol and Nicotine Use (GSCAN) GWAS of cigarettes per day to create the combined problematic tobacco use (PTU) phenotype [25]. The final GWAS summary statistics had an effective sample size of $n = 270,120$ individuals. We also report on model fit with just FTND as an indicator in the supplement (Alternative Models M2).

Cannabis use disorder (CUD). Summary statistics were derived from a GWAS meta-analysis [7] of DSM-IV and DSM-III-R cannabis abuse and dependence from the Psychiatric Genomics Consortium ($n = 5,289$ cases; $n = 10,004$ controls), ICD-10 cannabis use disorder from the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH) ($n = 2,758$ cases; $n = 53,326$ controls), and hospital-based diagnoses from deCODE ($n = 6,033$ cases; $n = 280,396$ controls). The final European-ancestry sample included 14,080 cases with CUD and 343,726 controls.

Opioid use disorder. Opioid use disorder (OUD): Summary statistics were derived from a meta-analysis [5] of GWASs of DSM-IV opioid abuse or dependence from Yale-Penn, and the Study of Addiction: Genetics and

Environment, and ICD-9/10 codes for opioid use disorder from the Million Veteran Program ($n = 10,544$ cases; $n = 72,163$ opioid-exposed controls).

Substance use summary statistics

Alcohol use. Alcohol use summary statistics were derived from the GSCAN GWAS [31] for current (this past week or average in the past year) reported drinks/week ($n = 537,349$). There was a strong correlation with lifetime PAU (SNP-rG between drinks/week and PAU = 0.77 ± 0.02) [21].

Lifetime tobacco use. Summary statistics came from the GSCAN GWAS of reported ever/never regular cigarette smoking (ever $n = 301,524$; never $n = 331,278$). There was a moderate correlation with PTU (SNP-rG = 0.28 ± 0.03).

Lifetime cannabis use. We used summary statistics from a meta-analysis of lifetime cannabis ever-use from the International Cannabis Consortium and UK Biobank (ever $n = 43,380$; never $n = 118,702$) [26]. There was a moderate correlation with CUD [7] (SNP-rG = 0.47 ± 0.05).

Stage-based behavioral constructs

The three-stage behavioral model of addiction focuses on "state" changes in substance use behaviors. Because GWASs measure individual differences in traits, we selected behaviors that (1) are known to convey vulnerability to each stage as proxies, and (2) are heritable.

Risk-taking and sensitivity to reward. A GWAS of risk-taking derived from a single item in the UK Biobank ("Would you describe yourself as someone who takes risks?"; data field #2040; risk-taker $n = 83,677$; non-risk taker $n = 244,662$) [32].

Executive function. The "preoccupation/anticipation" stage is characterized by maladaptive reward valuation and future planning. Recent work argues that this vulnerability is captured by executive functioning [33]. Summary statistics from a GWAS of a latent factor representing common executive functioning were used ($N = 427,037$) [22].

Negative emotionality and sensitivity to stress. The stage of withdrawal/ negative affect represents substance use functioning to mitigate aversive withdrawal symptoms, such as negative affect. Neuroticism has been found to modify stress sensitivity and neural reward processing [34]. Neuroticism was chosen as a trait-based measure representing liability to negative affect as opposed to depression because depression was included in the non-substance psychiatric disorder factor generation and because neuroticism includes trans-diagnostic constructs such as negative urgency (i.e., impulsive attempts to cope with negative affect) that may place individuals at risk for the negative reinforcing aspects of SUDs. We selected the largest GWAS of neuroticism as a heritable proxy ($N = 390,278$) [19].

Non-substance summary statistics

Summary statistics from the PGC Cross-disorder GWAS on the 8 disorders that were previously shown to fit a 3-factor confirmatory model were used [14]. These disorders included Schizophrenia [35], Bipolar Disorder [36], Major Depressive Disorder [37], Attention Deficit Hyperactivity Disorder [38], Obsessive Compulsive Disorder [39], Anorexia Nervosa [40], Tourette Syndrome [41], and Autism Spectrum Disorder [42] (See Supplemental Table 1 for details).

Statistical analysis

First, we estimated the pairwise genetic correlations between PAU, PTU, CUD, and OUD using Linkage Disequilibrium Score Regression (LDSR) [28]. After confirming that the four SUDs were significantly genetically correlated (see "Results"), we applied confirmatory factor analysis to the covariance matrix generated by LDSR using gSEM [43] with weighted least squares estimation; PAU, PTU, OUD, and CUD indicators were allowed to load freely on a single latent factor (i.e., *The Addiction-Risk-Factor*). The variance of this common latent factor was scaled to 1.0. A residual correlation between PAU and OUD was estimated as most of the data from both studies comes from the Million Veteran Project sample electronic health records (but see model fit without this residual correlation in the supplement—Alternative Models M1). In supplemental analyses, we also examined alternative two-factor models (Alternative Models M3).

Second, we used a series of structural regression models to estimate the extent to which genetic liability to stage-based constructs of addiction (i.e., risk-taking, executive function, and neuroticism) are related to *The Addiction-Risk-Factor*. Here, the *Addiction-Risk-Factor* variance was freed, and the OUD loading was set to 1.0 to scale the model. Intercorrelations were estimated between risk-taking, executive function, and neuroticism.

Third, we recreated the three factors from Lee et al. [14] (i.e., psychotic disorders, compulsive disorders, and early neuro-developmental disorders) and estimated their relationship with *The Addiction-Risk-Factor* while allowing for inter-factor correlations (the association between *The Addiction-Risk-Factor* and an alternative cross-disorder genetic model from a preprint [44] was also estimated; this alternative model is shown in Supplemental Fig. 7). This allowed us to estimate the unique association between each of the three psychopathology factors and *The Addiction-Risk-Factor* and to estimate variance that was residual to *The Addiction-Risk-Factor*. We then examined whether proxies for stage-based addiction constructs (i.e., risk-taking, executive function, and neuroticism) indirectly linked *The Addiction-Risk-Factor* to the three non-substance psychopathology factors using a multiple mediator model. We also conducted supplemental modified Q-Trait analyses [44] to examine the extent of the mediation (Supplemental Q-Trait Analysis). To estimate residual associations (i.e., direct paths) between the stage-based constructs and *The Addiction-Risk-Factor*, we re-structured the mediation model to one in which the three non-SUD psychopathology factors served as "mediators"
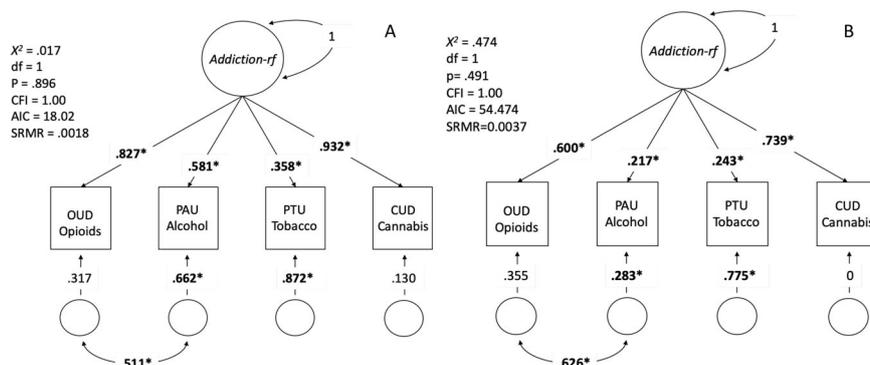

Fig. 1 Factor structure of 4 SUD GWAS. A The model, loadings, and fit for a model that allowed all four SUD categories to load on a latent factor. *A residual correlation between PAU and OUD was added to account for their assessment using electronic health records in the MVP cohort (models without residual correlations also fit well*: Supplemental Fig. 1). *Addiction-rf = The Addiction Risk-Factor.* **B** The same model, but accounting for common substance use (ever smoke, ever use marijuana, and drinks per week) as covariates at the indicator level, i.e., the three substance use measures are exogenous to all indicators in this model and the model represents the residual associations after accounting for substance use. Both models provided an excellent fit to the data. **Bold*** represents significance at $p < .05$. Note that in panel B, the residual of CUD is zero; this model constraint was necessary, as the model produced a negative residual without the constraint. Note: If you want to recreate the correlation matrix from both panels, the model with residual correlations cannot recover the implied correlation between PAU and OUD without taking the square root of the residual variance, rather than the value of the residual variance itself.

of the relationship between risk-taking, executive functioning, neuroticism, and *The Addiction-Risk-Factor*.

To separate the genetics of SUD from the genetics of substance use, we estimated models where substance *use* GWAS summary statistics were endogenous predictors of all measured variables in the model. For example, in the model estimating the association between *The Addiction-Risk-Factor* and psychiatric factors, the eight psychiatric disorders and the four SUD disorder variables were regressed on the three substance use variables. In this way, covariate effects were estimated simultaneously to our associations of interest.

RESULTS
The addiction risk factor
Genetic correlations between problematic alcohol use (PAU [21]), problematic tobacco use (PTU [24, 25]), cannabis use disorder (CUD [7]), and opioid use disorder (OUD [5]) ranged from 0.19 (S.E. = .04) to 0.78 (.09) (Supplemental Figs. 1 and 2). PTU showed the lowest SNP-rG with other SUD phenotypes [i.e., PAU = 0.19 (.04), CUD = 0.31 (.05), OUD = 0.26 (.08)] while OUD showed the highest [PAU = 0.69 (.07), CUD = 0.78 (.09)]. A confirmatory factor model specifying a unidimensional *Addiction-Risk-Factor* underlying the genetic covariance among PAU, PTU, CUD and OUD fit the data well [$X^2(1) = .017$, $p = .895$, CFI = 1, SRMR = .002; residual $r = .51$, $p = 0.016$; Fig. 1A]. Loadings were uniformly high except for PTU. Neither PAU nor PTU were impacted by the inclusion of non-diagnostic indices of addiction risk (Supplemental Results Alternative Models M1, M2; Supplemental Fig. 3). Alternative 2-factor models did not fit the data well (Alternative Models M3).

The inclusion of genetic liability to typical substance use did not modify the single factor structure of *The Addiction-Risk-Factor* (Fig. 1B); all SUDs continued to load significantly on the factor. However, factor loadings were lower for all substances, especially for PAU, which may be attributable to the high genetic correlation between drinks/week and PAU. Alternative parameterization of substance use as covariates did not improve model fit (Alternative Models M4).

Shared liability to stage-based behavioral phenotypes
Genetic liability to stage-based addiction constructs was shared with *The Addiction-Risk-Factor* (Fig. 2, Q-Trait Analysis in Supplemental Methods and Supplemental Fig. 4 for correlations). As expected, *The Addiction-Risk-Factor* was positively associated with genetic liability to risk-taking ($\beta = 0.45$) and neuroticism ($\beta = 0.25$), and negatively associated with executive function ($\beta = -0.17$; Fig. 2A). Despite significant genetic overlap between *The Addiction-Risk-Factor* and stage-based behavioral phenotypes, *The Addiction-Risk-Factor* retained unique variance (*Addiction-Risk-Factor* residual = 0.68). When conditioning for genetic liability for substance use, *The Addiction-Risk-Factor* remained significantly associated with increased genetic liability to risk-taking ($\beta = 0.22$) and neuroticism ($\beta = 0.18$) and decreased genetic liability to executive function ($\beta = -0.28$; Fig. 2B). Accounting for genetic liability for substance use substantially reduced the association between *The Addiction-Risk-Factor* and risk-taking from 0.45 to 0.22 ($p_{df = 1} = 4e{-}09$) and accentuated the negative association with executive function from $\beta = -0.17$ to -0.28 ($p_{(df = 1)} = 0.013$); there was a smaller effect on the association with neuroticism (from $\beta = 0.25$ to 0.18, $p_{(df = 1)} = 0.012$).

Shared liability to non-substance psychopathology
Genetic liability to non-substance psychopathology (i.e., compulsive disorders, psychotic disorders, and neurodevelopmental disorders) was shared with *The Addiction-Risk-Factor* (Fig. 3 with correlations in Supplemental Fig. 5, full models in Supplemental Fig. 6; Supplemental Fig. 7 shows results with an alternative cross-disorder model from a recent preprint [44]). Psychotic disorders ($\beta = 0.45$) and neurodevelopmental disorders ($\beta = 0.74$) were positively associated with *The Addiction-Risk-Factor* while compulsive disorders showed a negative association ($\beta = -0.32$; Fig. 3A). Due to the strong correlation between *The Addiction-Risk-Factor* and early-onset neurodevelopmental disorders (which includes ADHD) we allowed ADHD to load on *The Addiction-Risk-Factor* to
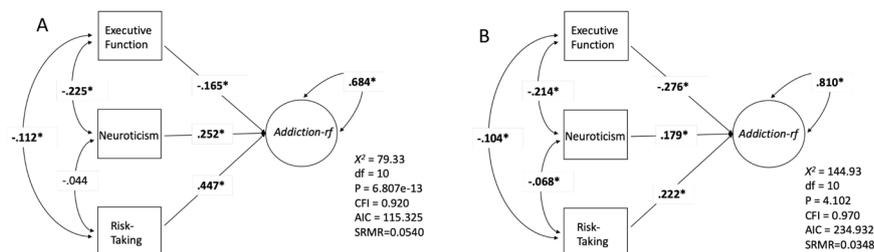

Fig. 2 Genetic associations between *The Addiction-Risk-Factor* and behavioral traits. Executive function, neuroticism, and risk-taking. **A** The model, fit, and regression pathways without accounting for common substance use. **B** Is the same model, but accounting for common substance use (ever smoke, ever use marijuana, and drinks per week) as covariates at the indicator level (regressed on all measured variables/ GWAS). **Bold*** represents significance at *p* < .05.
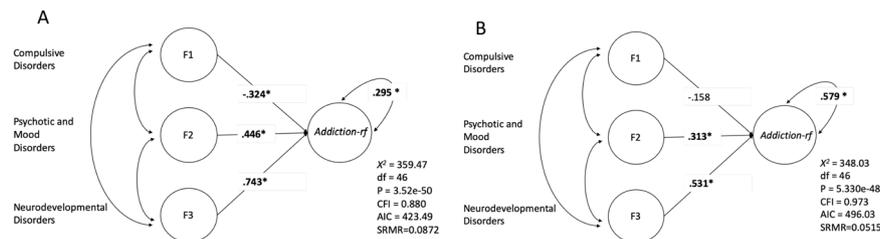

Fig. 3 Genetic associations between *The Addiction-Risk-Factor* and latent psychopathology factors. Compulsive disorders (F1; Tourette's syndrome, obsessive-compulsive disorder, and eating disorders), Psychotic Disorders (F2; Major Depressive Disorder, Schizophrenia, and Bipolar Disorder), and neurodevelopmental dysfunction (F3; ADHD, Autism, and Major Depressive Disorder). **A** The model, fit, and regression pathways without accounting for common substance use (model was scaled by setting the Opioid Use Disorder loading to 1). **B** Is the same model, but accounting for common substance use (ever smoke, ever use marijuana, and drinks per week) as covariates at the indicator level (regressed on all measured variables/GWAS), i.e., the three substance use measures are exogenous to all indicators in this model and the model is the residual associations after accounting for substance use. **Bold*** represents significance at *p* < .05. *Addiction-rf* = The Addiction-Risk-Factor.

Table 1. Behavioral liabilities mediate the association between *The Addiction-Risk-Factor* and psychiatric factors.

	F1 Indirect %	F1 Indirect P	F2 Indirect %	F2 Indirect P	F3 Indirect %	F3 Indirect P
No covariates						
Executive function	0.007	0.463	**0.051***	**0.001***	0.003	0.486
Neuroticism	0.005	0.842	0.003	0.836	0.013	0.777
Risk-taking	**−0.090***	**0.001***	**0.152***	**1.47E−08***	**0.043***	**0.013***
Controlling for genetics of substance use						
Executive function	0.013	0.233	**0.048***	**0.047***	−0.007	0.491
Neuroticism	0.011	0.687	−0.022	0.693	−0.040	0.689
Risk-taking	0.001	0.509	0.019	0.446	−0.002	0.741

Indirect associations from a mediation model (see Supplementary Fig. 8) where stage-based constructs link non-substance psychopathology (three factors from Lee et al.,) and *The Addiction-Risk-Factor*. F1 = compulsive disorders, F2 = Psychotic disorders, F3 = Neurodevelopmental disorders. The proportion accounted for by the indirect association (%) and the significance of the indirect association are shown. *P < .05 for the indirect association pathway.

control for ADHD; here, an association between *The Addiction-Risk-Factor* and early-onset neurodevelopmental disorders remained, but was significantly attenuated (from $\beta = 0.74$ to 0.43, $p_{(df = 1)} = $ 5e–5). When conditioning *The Addiction-Risk-Factor* for substance *use*, the psychotic and early neurodevelopmental disorder factors remained significantly associated with *The Addiction-Risk-Factor* (Fig. 3B). Despite the significant genetic overlap with other psychiatric disorder domains, *The Addiction-Risk-Factor* retained unique variance representing genetic liability specific to SUDs (*The Addiction-Risk-Factor* residual = 0.30, $p = 4.54e−3$). This unique variance remained significant when accounting for genetic liability to substance use (*The Addiction-Risk-Factor* residual = 0.58, $p = $ 0.015).

The specifications for the mediation models are shown in Supplemental Fig. 8 Genetic liability to risk-taking accounted for a proportion of the associations between all non-substance psychopathology domains and *The Addiction-Risk-Factor* (Table 1). Executive function uniquely indexed an indirect effect between psychotic disorders and *The Addiction-Risk-Factor* (Table 1). When conditioning *The Addiction-Risk-Factor* for genetic liability to substance use, risk-taking no longer accounted for a portion of the association between any non-substance psychopathology domain and *The Addiction-Risk-Factor*, but executive function continued to account for a proportion of the overlap (indirect effect of 0.048) between psychotic disorders and *The Addiction-Risk-Factor* (Table 1). Post hoc analyses revealed that executive function retained a unique association with *The Addiction-Risk-Factor* after accounting for genetic liability to both substance use and non-substance psychopathology (Supplementary Table 2).

DISCUSSION
We applied genomic structural equation modeling (gSEM) [23] to GWAS summary statistics to characterize the genetic influences shared across SUDs and estimate how common genetic liability is related to trait conceptualizations of a theoretical stage-based SUD model as well as to non-substance psychopathology. Three primary findings emerged. *First*, genetic risk for specific SUD phenotypes (i.e., PAU [21], PTU [24, 25], CUD [7], and OUD [5]) was largely attributable to a single Addiction risk factor, *The Addiction-Risk-Factor* (Fig. 1). *Second*, *The Addiction-Risk-Factor* was associated with genetic liability to trait representations of stage-based facets of addiction (risk-taking [binge/intoxication], executive function [preoccupation/anticipation], neuroticism [negative affect] [18]; Fig. 2). It was also associated with non-substance psychopathology factors (compulsive disorders, psychotic disorders, neurodevelopmental disorders; Fig. 3). Trait representations of stage-based facets of addiction partially accounted for the shared genetic liability between non-substance psychopathology and *The Addiction-Risk-Factor*. *Third*, associations between *The*

Addiction-Risk-Factor and stage-based constructs and non-substance psychopathology were largely independent of genetic liability to substance *use* phenotypes (i.e., tobacco use, cannabis use, alcoholic drinks/week). However, consistent with the stage-based model of addiction, accounting for substance use attenuated associations between risk-taking and *The Addiction-Risk-Factor* while potentiating associations with executive functioning. Collectively, our findings suggest that SUDs are characterized by a common genetic factor, *Addiction-Risk-Factor*.

The Addiction-Risk-Factor retains variance that is not shared with other psychopathology
After accounting for genetic liability to substance use, as well as the commonality between *The Addiction-Risk-Factor* and non-substance psychopathology, *The Addiction-Risk-Factor* retained significant variance. These data suggest that *The Addiction-Risk-Factor* may be characterized by unique pathways not shared with substance use or non-substance psychopathology, i.e., addiction is not the linear combination of substance use and psychopathology.

A single latent factor, fit these data well, but specific SUDs showed varying degrees of association. The illicit SUDs (CUD and OUD; Fig. 1) were almost entirely captured by the common latent factor. Notably, the loading for PTU on *The Addiction-Risk-Factor* was the smallest. One potential contributor to the residual variance of PTU may be the use of FTND and cigarettes/day as indices of PTU. Unlike the Diagnostic and Statistical Manual (DSM) criteria which index psychological and physiological aspects of tobacco use disorder, the FTND is an index of biochemical dependence and although used widely by investigators to define addiction, phenotypically shows only moderate agreement with DSM-defined nicotine dependence ($r = 0.50$; kappa = 0.3) [45].

Proxies of stage-based behavioral constructs and *The Addiction-Risk-Factor*
Behavioral stage-based models of SUD posit a cyclical relationship between positive reinforcement, negative reinforcement, and incentive salience [18] that we found can be (partially) captured by genetic liability to risk-taking, executive functioning, and negative emotionality (neuroticism). The strongest association with *The Addiction-Risk-Factor* was for risk-taking.

When substance use was included as a covariate in the model, the shared genetic loading between *The Addiction-Risk-Factor* and both risk-taking and neuroticism was attenuated down while the association with executive function increased. The reduction in the association with neuroticism is counter to expectations from the stage-based model which posits a more prominent role of negative affect for SUD relative to substance use. We speculate that neuroticism, which represents an amalgam of negative affect traits, may be too broad a construct when considering SUD-

specific negative affect; large-scale studies of domains of negative affectivity (e.g., negative urgency) are needed.

Non-substance psychopathology and *The Addiction-Risk-Factor*

We found that the three non-substance psychopathology clusters, derived from 8 psychiatric disorders [14], were genetically associated with *The Addiction-Risk-Factor*. The association with early neurodevelopmental disorders, which include ADHD, was the strongest. Cross-loading ADHD on *The Addiction-Risk-Factor* to condition on ADHD attenuated the loading but it remained high. Associations between *The Addiction-Risk-Factor* and the psychopathology clusters were greater than associations with trait representations of behavioral stages of addiction (with the exception of risk-taking). For instance, the genetic association between *The Addiction-Risk-Factor* and the two disorder clusters that included Major Depressive Disorder (i.e., psychotic disorders and early neurodevelopmental disorders) was greater in magnitude than the *Addiction-Risk-Factor*-neuroticism association. Interestingly the compulsive disorder factor did not show strong associations with *The Addiction-Risk-Factor*, suggesting that compulsive disorders and addiction-related compulsive behaviors have distinct etiologies.

Of the three behavioral correlates, risk-taking was the most prominent contributor to the association between *The Addiction-Risk-Factor* and all non-substance psychopathology factors. After accounting for substance use, only risk-taking and executive function mediated *The Addiction-Risk-Factor* associations with the psychotic disorder factor. Executive function maintained the only direct association with *The Addiction-Risk-Factor* after accounting for the genetics of substance use and genetics of non-substance psychopathology. Thus, we speculate that while risk-taking may characterize the genetic overlap between substance use and other psychopathology, executive function impairment is a risk factor that not only shapes the overlap between addiction and non-substance psychopathology but also explains variance in addiction above and beyond that overlap.

Limitations

There are several limitations. First, we had to restrict our analyses to individuals of European descent due to the lack of well-powered discovery GWAS informative for other ancestry groups. Second, to maximize the sample size of discovery GWASs, our alcohol and tobacco use GWAS incorporated measures of "problematic" use that, while genetically highly correlated with AUD and ND, may include behavioral patterns that are less severe than those represented by use disorder and were not assessed based on clinical presentation. Third, the analyses contain an over-representation of men, in part because the MVP samples contributed most of OUD and half of PAU and the MVP is ~90% male. Studies with larger numbers of women would allow stratified analyses to explore the differences between sexes observed in epidemiological studies. Fourth, while it is unlikely that individuals completed assessments of risk-taking, neuroticism, and executive function while under the influence of substances, how substance use may have influenced these assessments cannot be determined. Fifth, though significant, mediation pathways were small in effect. Sixth, how these processes effect phenotypic patterns is unknown, however, twin studies support a common factor model as well [4].

CONCLUSIONS

Common genetic liability undergirds distinct SUDs and shares variance with putative behavioral intermediary phenotypes/SUD risk factors and non-substance psychopathology. This addiction genetic factor is more than a linear combination of substance use and psychopathology; it represents a unique addiction dimension that is partially captured by executive functions.

REFERENCES

1. Degenhardt L, Charlson F, Ferrari A, Santomauro D, Erskine H, Mantilla-Herrara A, et al. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Psychiatry. 2018;5:987–1012.
2. Bhalla IP, Stefanovics EA, Rosenheck RA. Clinical epidemiology of single versus multiple substance use disorders. Med Care. 2017;55:S24–S32.
3. Merikangas KR, Kalaydjian A. Magnitude and impact of comorbidity of mental disorders from epidemiologic surveys. Curr Opin Psychiatry. 2007;20:353–8.
4. Palmer RH, Button TM, Rhee SH, Corley RP, Young SE, Stallings MC, et al. Genetic etiology of the common liability to drug dependence: evidence of common and specific mechanisms for DSM-IV dependence symptoms. Drug Alcohol Depend. 2012;123 Suppl:S24–32.
5. Zhou H, Rentsch CT, Cheng Z, Kember RL, Nunez YZ, Sherva RM, et al. Association of OPRM1 functional coding variant with opioid use disorder: a genome-wide association study. JAMA Psychiatry. 2020. https://doi.org/10.1001/jamapsychiatry.2020.1206.
6. Sanchez-Roige S, Palmer AA, Clarke TK. Recent efforts to dissect the genetic basis of alcohol use and abuse. Biol Psychiatry. 2020;87:609–18.
7. Johnson EC, Demontis D, Thorgeirsson TE, Walters RK, Polimanti R, Hatoum AS, et al. A large-scale genome-wide association study meta-analysis of cannabis use disorder. Lancet Psychiatry. 2020;0:1032–45.
8. Polimanti R, Walters RK, Johnson EC, McClintick JN, Adkins AE, Adkins DE, et al. Leveraging genome-wide data to investigate differences between opioid use vs. opioid dependence in 41,176 individuals from the Psychiatric Genomics Consortium. Mol Psychiatry. 2020;25:1673–87.
9. Kranzler HR, Zhou H, Kember RL, Vickers Smith R, Justice AC, et al. Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. Nat Commun. 2019;10:1499.
10. Sanchez-Roige S, Palmer AA, Fontanillas P, Elson SL, Research Team, the Substance Use Disorder Working Group of the Psychiatric Genomics C., Adams MJ, et al. Genome-wide association study meta-analysis of the alcohol use disorders identification test (AUDIT) in two population-based cohorts. Am J Psychiatry. 2019;176:107–18.
11. Walters RK, Polimanti R, Johnson EC, McClintick JN, Adams MJ, Adkins AE, et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. Nat Neurosci. 2018;21:1656–69.
12. Mallard TT, Savage JE, Johnson EC, Huang Y, Edwards AC, Hottenga JJ, et al. Item-level genome-wide association study of the alcohol use disorders identification test in three population-based cohorts. Am J Psychiatry. 2021. https://doi.org/10.1176/APPI.AJP.2020.20091390.
13. Sanchez-Roige S, Cox NJ, Johnson EO, Hancock DB, Davis LK. Alcohol and cigarette smoking consumption as genetic proxies for alcohol misuse and nicotine dependence. Drug Alcohol Depend. 2021;221:108612.
14. Lee PH, Anttila V, Won H, Feng Y-CA, Rosenthal J, Zhu Z, et al. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. Cell. 2019;179:1469–1482.e11.
15. Carey CE, Agrawal A, Bucholz KK, Hartz SM, Lynskey MT, Nelson EC, et al. Associations between polygenic risk for psychiatric disorders and substance involvement. Front Genet. 2016;7:149.
16. Karlsson Linnér R, Mallard TT, Barr PB, Sanchez-Roige S, Madole JW, et al. Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. Nat. Neurosci. 2021. https://doi.org/10.1038/s41593-021-00908-3.
17. Abdellaoui A, Smit DJA, van den Brink W, Denys D, Verweij KJH. Genomic relationships across psychiatric disorders including substance use disorders. Drug Alcohol Depend. 2021;220:108535.
18. Goldstein RZ, Volkow ND. Drug addiction and its underlying neurobiological basis: neuroimaging evidence for the involvement of the frontal cortex. Am J Psychiatry. 2002;159:1642–52.
19. Nagel M, Watanabe K, Stringer S, Posthuma D, van der Sluis S. Item-level analyses reveal genetic heterogeneity in neuroticism. Nat Commun. 2018;9:905.

20. Koob GF, Volkow ND. Neurobiology of addiction: a neurocircuitry analysis. Lancet Psychiatry. 2016;3:760–73.

21. Zhou H, Sealock JM, Sanchez-Roige S, Clarke TK, Levey DF, Cheng Z, et al. Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. Nat Neurosci. 2020. https://doi.org/10.1038/s41593-020-0643-5.

22. Hatoum A, Mitchell E, Morrison C, Evans L, Keller M, Friedman N, et al. GWAS of over 427,000 individuals establishes GABAergic and synaptic molecular pathways as key for cognitive executive functions. Preprint at *bioRxiv* (2019). https://doi.org/10.1101/674515.

23. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav. 2019;3:513–25.

24. Hancock DB, Guo Y, Reginsson GW, Gaddis NC, Lutz SM, Sherva R, et al. Genome-wide association study across European and African American ancestries identifies a SNP in DNMT3B contributing to nicotine dependence. Mol Psychiatry. 2018;23:1–9.

25. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet. 2019;51:237–44.

26. Pasman JA, Verweij K, Gerring Z, Stringer S, Sanchez-Roige S, Treur JL, et al. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. Nat Neurosci. 2018;21:1161–70.

27. Edenberg HJ, McClintick JN. Alcohol dehydrogenases, aldehyde dehydrogenases, and alcohol use disorders: a critical review. Alcohol: Clin Exp Res. 2018;42:2281–97.

28. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47:291–5.

29. Zhou H, Sealock JM, Sanchez-Roige S, Clarke TK, Levey DF, Cheng Z, et al. Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. Nat Neurosci. 2020;23:809–18.

30. Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA, et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nat Genet. 2018;50:229–37.

31. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat Genet. 2019;51:237–44.

32. Strawbridge RJ, Ward J, Lyall LM, Tunbridge EM, Cullen B, Graham N, et al. Genetics of self-reported risk-taking behaviour, trans-ethnic consistency and relevance to brain gene expression. Transl Psychiatry. 2018;8:178.

33. Volkow ND, Koob GF, McLellan AT. Neurobiologic advances from the brain disease model of addiction. N Engl J Med. 2016;374:363–71.

34. Bondy E, Baranger DA, Balbona JV, Sputo K, Paul SE, Oltmanns T, et al. Neuroticism and reward-related ventral striatum activity: probing vulnerability to stress-related depression. Preprint at *bioRxiv* (2020). https://doi.org/10.31234/osf.io/5wd3k.

35. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511:421–7.

36. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. Nat Genet 2019;51:793–803.

37. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci. 2019;22:343–52.

38. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet. 2019;51:63–75.

39. Arnold PD, Askland KD, Barlassina C, Bellodi L, Bienvenu OJ, Black D, et al. Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. Mol Psychiatry. 2018;23:1181–8.

40. Watson HJ, Yilmaz Z, Thornton LM, Hübel C, Coleman J, Gaspar HA, et al. Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. Nat Genet. 2019;51:1207–14.

41. Yu D, Sul JH, Tsetsos F, Nawaz MS, Huang AY, Zelaya I, et al. Interrogating the genetic determinants of Tourette's syndrome and other tiC disorders through genome-wide association studies. Am J Psychiatry. 2019;176:217–27.

42. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, et al. Identification of common genetic risk variants for autism spectrum disorder. Nat Genet. 2019;51:431–44.

43. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. Genomic SEM provides insights into the multivariate genetic architecture of complex traits. Preprint at *bioRxiv* (2018). https://doi.org/10.1101/305029.

44. Grotzinger AD, Mallard TT, Akingbuwa WA, Ip HF, Adams MJ, Lewis CM, et al. Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic, and molecular genetic levels of analysis. medRxiv. 2020;18:18–19.

45. Agrawal A, Scherrer JF, Pergadia ML, Lynskey MT, Madden PA, Sartor CE, et al. A latent class analysis of DSM-IV and fagerström (FTND) criteria for nicotine dependence. Nicotine Tob Res. 2011;13:972–81.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

ASH conducted the analysis and wrote the manuscript. ECJ, SMC assisted with the analysis and creation of the figures. RP, HZ, RW, JG, HE, RB, and AA, provided critical commentary on the manuscript and argumentation. RB and AA oversaw all analysis. ASH had access to all data and takes responsibility for the analysis herein.

ADDITIONAL INFORMATION