Psychological Medicine

cambridge.org/psm

Invited Review

[‡]The online version of this article has been updated since original publication. A notice detailing the changes has also been published at: https://doi.org/10.1017/S0033291722000629

Cite this article: Deak JD, Johnson EC (2021). Genetics of substance use disorders: a review. *Psychological Medicine* **51**, 2189–2200. https://doi.org/10.1017/S0033291721000969

Received: 2 October 2020
Revised: 22 February 2021
Accepted: 4 March 2021
First published online: 21 April 2021

Key words:
substance use disorders; genetics; genome wide association study; twin and family studies; heritability; genetic epidemiology

Author for correspondence:
Emma C. Johnson,
E-mail: emma.c.johnson@wustl.edu

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

CAMBRIDGE UNIVERSITY PRESS

Genetics of substance use disorders: a review[‡]

Joseph D. Deak[1,2] and Emma C. Johnson[3]

[1]Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA; [2]Department of Psychiatry, Veterans Affairs Connecticut Healthcare Center, West Haven, CT, USA and [3]Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA

Abstract

Substance use disorders (SUDs) are prevalent and result in an array of negative consequences. They are influenced by genetic factors (h^2 = ~50%). Recent years have brought substantial progress in our understanding of the genetic etiology of SUDs and related traits. The present review covers the current state of the field for SUD genetics, including the epidemiology and genetic epidemiology of SUDs, findings from the first-generation of SUD genome-wide association studies (GWAS), cautions about translating GWAS findings to clinical settings, and suggested prioritizations for the next wave of SUD genetics efforts. Recent advances in SUD genetics have been facilitated by the assembly of large GWAS samples, and the development of state-of-the-art methods modeling the aggregate effect of genome-wide variation. These advances have confirmed that SUDs are highly polygenic with many variants across the genome conferring risk, the vast majority of which are of small effect. Downstream analyses have enabled finer resolution of the genetic architecture of SUDs and revealed insights into their genetic relationship with other psychiatric disorders. Recent efforts have also prioritized a closer examination of GWAS findings that have suggested non-uniform genetic influences across measures of substance use (e.g. consumption) and problematic use (e.g. SUD). Additional highlights from recent SUD GWAS include the robust confirmation of loci in alcohol metabolizing genes (e.g. *ADH1B* and *ALDH2*) affecting alcohol-related traits, and loci within the *CHRNA5-CHRNA3-CHRNB4* gene cluster influencing nicotine-related traits. Similar successes are expected for cannabis, opioid, and cocaine use disorders as sample sizes approach those assembled for alcohol and nicotine.

Introduction

Substance use disorders (SUDs) are heritable psychiatric disorders [Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5); American Psychiatric Association, 2013] that are influenced by both environmental and genetic factors. Given the public health burden of SUDs, a better understanding of SUD etiology is of wide-reaching importance. Genetic studies have begun to elucidate the molecular mechanisms underlying SUDs and related traits, including other psychiatric conditions with which SUDs frequently co-occur (Grant et al., 2016; Kessler, 2004). While there are challenges inherent in studying complex, polygenic traits such as SUDs, it is hoped that better understanding the genetic basis of risk for developing SUDs will eventually help inform SUD prevention and treatment. In this review, we cover SUD epidemiology, conclusions from twin and family studies of SUDs, and findings from more recent molecular genetic studies[1]; finally, we summarize the current state of the field and suggest future directions.

Definition of SUD

SUDs are defined by the DSM-5 (American Psychiatric Association, 2013) as the presence of at least two of 11 criteria in a 12-month period, with disorder severity indexed by the number of criteria endorsed (2–3 = mild; 4–5 = moderate; ⩾6 = severe). Broadly speaking, DSM-5 SUD criteria correspond to the presence of substance-related problems, such as increased use, unsuccessful attempts to stop or cut down, continued use despite negative physical, psychological, and social consequences, persistent craving, development of tolerance, and symptoms of withdrawal. These criteria can be assessed in relation to multiple substances (e.g. alcohol, nicotine, cannabis, opioid, cocaine).

Prior to DSM-5, DSM-IV distinguished substance abuse from dependence. Substance abuse required the endorsement of at least one of four abuse criteria, while substance dependence required at least three of seven dependence criteria. Research favored a unidimensional diagnosis over the separation of abuse and dependence (Hasin et al., 2013). With the revised DSM-5 SUD criteria, the substance abuse 'legal problems' criterion was removed and a craving criterion was added. Compton, Dawson, Goldstein, and Grant (2013) found that a threshold of

⩾4 DSM-5 criteria (i.e. moderate severity) demonstrated optimal correspondence with DSM-IV dependence for alcohol, cocaine, and opioid use disorders.

The present review will focus on five SUDs that have demonstrated substantial progress in molecular genetics in recent years: alcohol use disorder (AUD), nicotine use disorder (NicUD), cannabis use disorder (CanUD), opioid use disorder (OUD), and cocaine use disorder (CocUD). Many studies reviewed utilized DSM-IV defined substance abuse and dependence.

Epidemiology

SUDs are highly prevalent behaviors associated with an array of negative outcomes. Epidemiological estimates suggest that up to 29.1% and 27.9% of individuals will meet the criteria for AUD and NicUD, respectively, in their lifetime, with lower lifetime prevalence rates for CanUD (6.3%), OUD (2.1%), and CocUD (2.4%) (Grant et al., 2016). All SUDs contribute to increased rates of injuries, elevated risk of other disorders, and pose a large economic cost worldwide (Degenhardt & Hall, 2012). For example, AUD and NicUD, respectively, contribute to 3 million (5.3%) and 7 million (12.3%) worldwide deaths annually, making both among the leading causes of global mortality (Global Status Report on Alcohol and Health, 2018; WHO Report on the Global Tobacco Epidemic, 2017). SUDs that occur at lower rates (e.g. OUD) also have severe impact; the USA is currently combating an opioid use public health crisis, with an estimated 47 600 individuals dying from opioid overdoses in 2018 (Hedegaard, Miniño, & Warner, 2020).

As reviewed by Koob and Le Moal (2001), the cycle of addictions (including SUDs) can be thought of as having three main components: preoccupation-anticipation, binge-intoxication, and withdrawal-negative affect. This model aligns with DSM criteria, while allowing for the influence of genetic vulnerabilities and environmental risk and protective factors at different stages. New research also suggests this stage-based model is supported by genomic data (Hatoum et al., 2021). Family history of SUDs, peer substance use, lower socioeconomic status (SES), and psychiatric comorbidities are all associated with increased risk of developing a SUD (Stone, Becker, Huber, & Catalano, 2012). Some of these are potential *consequences* of SUDs as well as risk factors (e.g. socioeconomic hardship, poorer mental health prognosis), and contribute to the negative outcomes associated with SUDs (Kendler, Ohlsson, Karriker-Jaffe, Sundquist, & Sundquist, 2017).

Genetic epidemiology

Twin and family studies have demonstrated strong familial inheritance patterns for SUDs (Prom-Wormley, Ebejer, Dick, & Bowers, 2017). Heritability (h^2) estimates across SUDs vary, but broadly suggest that genetic influences account for approximately 50% of the risk. Quantitative genetic studies have also suggested that, in addition to the presence of substance-specific influences for SUDs – with nicotine and opiates showing the most evidence of substance-specific genetic factors (Kendler, Myers, & Prescott, 2007; Tsuang et al., 1998) – there are heritable factors that contribute to SUDs more broadly (Kendler et al., 2007; Kendler, Prescott, Myers, & Neale, 2003).

Alcohol use disorder

Heritability estimates for AUD range from ~0.50 to 0.64 (Heath et al., 1997; Kendler, 2001), with a recent meta-analysis reporting h^2 of ~0.50 (Verhulst, Neale, & Kendler, 2015). Heritability estimates for AUD diagnosis tend to be slightly higher than for other alcohol-related traits, such as alcohol use initiation ($h^2 = 0.30$–0.40; Koopmans, Slutske, Van Baal, & Boomsma, 1999) and alcohol use frequency ($h^2 = 0.37$–0.47; Viken, Kaprio, Koskenvuo, & Rose, 1999), which is consistent with prior twin studies suggesting that environmental influences may have a more pronounced impact on initiation, while genetic factors are more influential in progression to heavier use and substance-related problems (Kendler, Karkowski, Neale, & Prescott, 2000).

Nicotine use disorder

Heritable factors contribute across the stages of cigarette smoking and NicUD, with a range of heritability estimates for nicotine dependence (ND) between ~0.30 and 0.70 (Agrawal et al., 2012; Sullivan & Kendler, 1999). Variability in reported h^2 results for NicUD could, at least in part, be due to the different ways in which NicUD-related problems have been assessed [e.g. Fagerström Tolerance Questionnaire (FTQ), Fagerström Test for Nicotine Dependence (FTND)] in comparison to NicUD as determined by DSM diagnostic criteria (Cohen, Myers, & Kelly, 2002; Payne, Smith, McCracken, McSherry, & Antony, 1994).

Cannabis use disorder

Heritability estimates from twin studies of CanUD range from ~0.51 to 0.59, slightly higher than the estimates for cannabis use/initiation (~0.40–0.48; Agrawal & Lynskey, 2006; Verweij et al., 2010). Twin and family studies have found shared genetic and environmental influences across the stages of cannabis use and abuse (Agrawal, Neale, Jacobson, Prescott, & Kendler, 2005; Van den Bree, Johnson, Neale, & Pickens, 1998). Gillespie, Neale, and Kendler (2009) explored this further, finding that *availability* of cannabis explained nearly all of the shared environmental variance in cannabis initiation and abuse, initiation mediated the influence of availability on abuse, and a large proportion of the genetic variance in abuse (62%) was shared with initiation.

Opioid use disorder

Twin and family studies have estimated that ~50% of the liability to opioid dependence is due to additive genetic factors (Berrettini, 2017; Kendler, Jacobson, Prescott, & Neale, 2003; Tsuang, Bar, Harley, & Lyons, 2001). Tsuang et al. (1998) estimated that 38% of the variation in opioid addiction was due to genetic factors *specific* to opioids (i.e. not shared with other substances).

Cocaine use disorder

Estimates of the heritability of CocUD range from ~0.40 to 0.80, with evidence of a common genetic vulnerability with other SUDs, especially cannabis, and little evidence of cocaine-specific genetic influences (Kendler et al., 2007).

Genetic correlations amongst SUDs

Twin studies have also been used to assess the genetic correlations (r_g) amongst substance use and SUDs. Population-based twin estimates of r_g may be less prone to biases inherent in modern genome-wide studies that rely on large biobanks with

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

unrepresentative sample characteristics (e.g. the relatively healthy and high SES makeup of the UK Biobank). Briefly, findings from twin and family studies suggest common genetic factors shared amongst substance use ($r_g = 0.14–0.31$ for alcohol, tobacco, and cannabis), with stronger estimates of shared genetic overlap amongst measures of *problem* use ($r_g = 0.56–0.62$; Young, Rhee, Stallings, Corley, & Hewitt, 2006).

Genetics of SUDs

Genome-wide association studies (GWAS) of SUDs have rapidly increased in both sample size and locus identification during the past 5 years (reviewed in Hancock, Markunas, Bierut, & Johnson, 2018; Johnson, Chang, & Agrawal, 2020*a*). Still, the sample sizes for SUD GWAS have lagged behind those of lifetime ever-use or consumption [e.g. drinks per week (DPW)], largely due to the additional burden associated with administering and undergoing a comprehensive SUD assessment in comparison to more readily administered survey and screener questionnaires. In addition, there are a variety of considerations in the recruitment of individuals with SUDs (Fisher & Jaber, 2019), including that these individuals may be more difficult to reach and/or less willing to participate in research studies. Sample diversity has also been limited; the majority of SUD GWAS to date are primarily composed of individuals of European ancestry and findings may not generalize to those of other ancestries.

One overarching question that has emerged from the first-generation of well-powered SUD GWAS is whether measures of non-problematic substance use have divergent genetic underpinnings from SUDs, and if so, to what extent. Another area of interest has been dissecting the genetic relationships between SUDs, other psychiatric disorders, and relevant complex traits; by leveraging large GWAS and advanced statistical genetics methods [e.g. cross-trait genetic correlations, genomicSEM (Grotzinger et al., 2019)], interesting patterns of pleiotropy have emerged (Abdellaoui, Smit, Van Den Brink, Denys, & Verweij, 2021; Hatoum et al., 2021; Jang et al., 2020). Notable SUD GWAS loci are summarized in Table 1.

Alcohol use disorder

Up until the past 5 years, there was limited progress in the identification of replicable genetic loci associated with AUD, excepting the well-established influence of genes encoding alcohol metabolism enzymes [e.g. alcohol dehydrogenase 1B (*ADH1B*), aldehyde dehydrogenase 2 (*ALDH2*); reviewed in Edenberg & Mcclintick, 2018]. Recent studies have demonstrated that, similar to other complex traits, larger sample sizes have aided the successful detection of loci influencing AUD and alcohol-related outcomes (reviewed in Deak, Miller, & Gizer, 2019; Sanchez-Roige, Palmer, & Clarke, 2020). These have replicated genome-wide significant (GWS) associations for loci in the *ADH1B* gene (e.g. rs1229984 and rs2066702) with AUD (Gelernter et al., 2014*a*; Kranzler et al., 2019; Walters et al., 2018; Zhou et al., 2020*a*, *b*) and with various measures of alcohol use and consumption (Clarke et al., 2017; Gelernter et al., 2019; Kranzler et al., 2019; Liu et al., 2019; Sanchez-Roige et al., 2019*a*, *b*; Xu et al., 2015). Similar success has been found for loci mapped to *ALDH2* (i.e. rs671) with robust associations found with alcohol dependence and alcohol-related traits (i.e. maximum drinks, flushing response) in East-Asian and Thai populations (Gelernter et al., 2018; Li, Zhao, & Gelernter, 2012; Quillen et al., 2014), and for alcohol drinking status in East-Asian populations (Jorgenson et al., 2017).

Associations with other loci have begun to be consistently identified as well. Recent GWAS have robustly identified associations between genetic variation in *DRD2* (Dopaminergic Receptor D2) and AUD (rs4936277, rs61902812; Kranzler et al., 2019) and problematic alcohol use (PAU; rs138084129, rs6589386; Zhou et al. 2020*a*, *b*), as well as gene-based associations with alcohol problems, as indexed by Alcohol Use Disorders Identification Test (AUDIT) scores (Sanchez-Roige et al., 2019*a*, *b*). Other loci, including the *GCKR* gene (Glucokinase Regulator; lead SNP: rs1260326), are associated with AUD and alcohol use problems (Kranzler et al., 2019; Sanchez-Roige et al., 2019*a*, *b*; Zhou et al., 2020*a*, *b*), as well as consumption (Clarke et al., 2017; Kranzler et al., 2019; Liu et al., 2019; Sanchez-Roige et al., 2019*a*, *b*). A *KLB* (Klotho Beta) variant was also found to be associated with PAU (rs13129401; Zhou et al., 2020*a*, *b*), and with AUDIT measures of alcohol problems (AUDIT-P; Sanchez-Roige et al., 2019*a*, *b*) and consumption (AUDIT-C; Kranzler et al., 2019; Sanchez-Roige et al., 2019*a*, *b*). There also have been associations between variants in the *SLC39A8* gene and AUD (rs13107325; Kranzler et al., 2019), AUDIT-P (rs13135092; Sanchez-Roige et al., 2019*a*, *b*) and AUDIT-C [rs13107325 (Kranzler et al., 2019); rs13135092 (Sanchez-Roige et al., 2019*a*, *b*)].

Methodological advancements (e.g. LD score regression, Bulik-Sullivan et al., 2015) modeling aggregate genetic risk across the genome have furthered our understanding of the genetic architecture of AUD. LD score regression approaches have generated SNP-heritability (h^2_{SNP}) estimates of ~0.07–0.10 for measures of PAU (Zhou et al., 2020*a*, *b*), and demonstrated positive genetic correlations between PAU and smoking cigarettes regularly, lifetime cannabis use, major depression, and risk-taking (Table 1; Zhou et al., 2020*a*, *b*). Additionally, varying patterns of genetic correlations suggest only a partial genetic overlap between AUD and alcohol consumption (Sanchez-Roige et al., 2020), likely due at least in part to differences in levels of drinking and overarching pathology across samples. A recent study reported a stronger correlation between DPW and PAU ($r_g = 0.77$; Zhou et al., 2020*a*, *b*), highlighting that high consumption is a necessary component of AUD; however, similar to other measures of alcohol consumption (e.g. AUDIT-C), DPW has demonstrated negligible genetic correlations ($r_g = -0.02$ to 0.08) with other psychiatric disorders [attention-deficit hyperactivity disorder (ADHD), bipolar disorder, major depression, schizophrenia; Jang et al., 2020], while PAU has shown stronger genetic overlap with these disorders ($r_g = 0.32–0.39$; Zhou et al., 2020*a*, *b*). These most recent results from Jang et al. (2020) and Zhou et al. (2020*a*, *b*), respectively, suggest that while DPW is a component of problematic use, PAU seems to capture shared genetic risk with other psychiatric disorders, while DPW does not.

Despite divergent patterns of genetic overlap suggesting non-uniform genetic influences, it should be noted that genes influencing alcohol-metabolizing enzymes (e.g. *ADH1B*, *ALDH2*) directly affect alcohol consumption, and in turn, play a role in the risk of AUD development. The coding variants in these genes provide a protective effect for AUD by producing aversive effects when drinking alcohol, often resulting in lower levels of consumption and AUD risk (Edenberg & Mcclintick, 2018). However, it is likely that thousands of additional genetic loci play a role beyond the genes encoding alcohol metabolizing enzymes. Additional studies have examined subdomains of alcohol consumption, suggesting potential etiological differences

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

Table 1. Summary of epidemiology, genetic epidemiology, and molecular genetic findings for substance use disorders

	Alcohol use disorder (AUD)	Nicotine use disorder (NicUD)	Cannabis use disorder (CanUD)	Opioid use disorder (OUD)	Cocaine use disorder (CocUD)[a]
SUD epidemiology and recent developments	∘ Lifetime prevalence rate = 29.1%[41] ∘ Alcohol use and intoxication contributes to three million worldwide deaths annually[122]	∘ Lifetime prevalence rate = 27.9%[41] ∘ Nicotine use and related-disease contributes to seven million worldwide deaths annually[123]	∘ Lifetime prevalence rate = 6.3%[41] ∘ Recent legalization in Western countries is correlated with increased use, including among pregnant women; it remains to be seen whether this influences prevalence rates of CUD[15]	∘ Lifetime prevalence rate = 2.1%[41] ∘ Despite lower prevalence relative to other SUDs, OUD poses large disease burden due to overdose deaths[53]	∘ Lifetime prevalence rate = 2.4%[41] ∘ From 2012 to 2018, the rate of overdose deaths related to cocaine use increased from 1.4% to 4.5%[53]
SUD genetic epidemiology	∘ AUD heritability (h^2) = 0.50–0.64[51,61,117] ∘ SNP-heritability (h^2_{SNP}) = ∼0.07–0.10[128]	∘ NicUD heritability (h^2) = ∼0.30–0.70[4,111] ∘ SNP-heritability (h^2_{SNP}) = ∼0.09[98]	∘ CanUD heritability (h^2) = 0.40–0.80[118] ∘ SNP-heritability (h^2_{SNP}) = ∼0.07–0.12[61]	∘ OUD heritability (h^2) = ∼0.50[8,67,114] ∘ SNP-heritability (h^2_{SNP}) = ∼0.11[127]	∘ CocUD heritability (h^2) = 0.40–0.80[69] ∘ SNP-heritability (h^2_{SNP}) = ∼0.27–0.30[14,55]
Notable GWAS risk-loci to date **BOLD** = loci have achieved $p \leq 5.0\text{x}10^{-9}$ with respective SUD in at least one study	∘ ***ADH1B***[b,c] (rs1229984)[17,32,36,78,84,105,121,124,128] ∘ ***ALDH2***[b,c] (rs671)[37,62,81,99] ∘ ***DRD2***[b] (e.g. rs4936277)[78,105,128] ∘ ***KLB***[b,c] (e.g. rs13129401)[78,105,128] ∘ ***GCKR***[b,c] (rs1260326)[17,78,84,105,128] ∘ ***SLC39A8***[b,c] (rs13107325)[78,84,105]	∘ ***CHRNA5***[b,c] (rs16969968)[31,45,46,84,98] ∘ ***CHRNA5-A3-B4***[b,c] (multiple loci)[31,45,46,84,98] ∘ ***CHRNA4***[b,c] (rs151176846)[45,46,84,98] ∘ *DNMT3B*[b] (rs910083)[46] ∘ *DBH*[b,c] (rs13284520)[84,98]	∘ ***FOXP2***[b] (rs7783012)[61] ∘ ***CHRNA2***[b] (rs4732724)[23,61] ∘ *EPHX2*[b] (rs4732724)[61] ∘ *CSMD1*[b] (rs77378271)[107] ∘ *PDE4B*[b] (gene-wise)[61]	∘ ***OPRM1***[b] (rs1799971)[127] ∘ ***CNIH3***[b] (rs10799590)[92] ∘ ***KCNG2***[b] (rs62103177)[33] ∘ *RGMA*[b] (rs12442183)[16] ∘ *BEND4*[c] (gene-wise)[96]	∘ *FAM53B*[b] (rs2629540)[34] ∘ *HIST1H2BD*[b] (gene-wise)[14] ∘ *C1QL2*[b] (gene-wise)[55] ∘ *STK38*[b] (gene-wise)[55] ∘ *KCTD20*[b] (gene-wise)[55]
Notable genetic correlations (r_g) with psychiatric/substance use traits	∘ Drinks per week (r_g = +0.77)[128] ∘ Ever smoked regularly (r_g = +0.55)[128] ∘ Lifetime cannabis use (r_g = +0.39) [128] ∘ Major depression (r_g = +0.39)[128] ∘ Risk-taking (r_g = +0.30)[128]	∘ Alcohol dependence (r_g = +0.56)[98] ∘ Cigarettes per day (r_g = +0.95)[98] ∘ Major depression (r_g = +0.38)[98] ∘ Schizophrenia (r_g = +0.16)[98] ∘ Smoking initiation (r_g = +0.40)[98]	∘ Alcohol use disorder (r_g = +0.55)[61] ∘ Educational attainment (r_g = −0.39)[61] ∘ Lifetime cannabis use (r_g = +0.50)[61] ∘ Schizophrenia (r_g = +0.31)[61] ∘ Smoking initiation (r_g = +0.66)[61]	∘ ADHD (r_g = +0.36)[127] ∘ Alcohol dependence (r_g = +0.73)[127] ∘ Drinks per week (r_g = +0.38)[127] ∘ Ever smoked regularly (r_g = +0.51)[127] ∘ Major depression (r_g = +0.35)[127]	∘ ADHD (r_g = +0.50)[14] ∘ Ever smoked regularly (r_g = +0.34)[14] ∘ Major depression (r_g = +0.40)[14] ∘ Risk-taking (r_g = +0.35)[14] ∘ Schizophrenia (r_g = +0.20)[14]
Notable CNV and exome/genome sequencing efforts	Genome-wide meta-analysis of CNV associations in AUD cases[112]: ∘ identified nine CNV regions suggestively associated with AUD (e.g. 5q21.3 deletion)	Exome-chip meta-analysis fine mapped rare coding variants for nicotine use outcomes[10]: ∘ identified 124 significant associations. ∘ 1.0–2.2% of phenotypic variance explained by rare variation	Low-coverage WGS found two gene regions significantly associated with CanUD[39]: ∘ *C1orf110* gene (protein-coding region) ∘ *MEF2B* gene (regulatory region)	Three significantly associated CNVs[82]: ∘ a 18q12.3 deletion ∘ a Xq28 deletion ∘ a 18q12.3 deletion	Several targeted sequencing[44] and CNV studies (e.g. *NSF* gene)[13] have been conducted, although no strong evidence has emerged. CocUD whole-genome, whole-exome, and CNV studies are needed
Notable efforts incorporating non-European populations	∘ Kranzler et al. (2019)[78]: AUD cases: EUR = 34 658; AFR = 17 267; LAT = 3449; EAA = 164; SAA = 44 ∘ Walters et al. (2018)[121]: AUD cases: EUR = 11 569; AFR = 3335	∘ Quach et al. (2020)[98]: *N* = 46 213 EUR smokers; *N* = 11 787 AFR smokers ∘ Hancock et al. (2018*a*)[46]: *N* = 28 677 EUR smokers; *N* = 9925 AFR smokers	∘ Johnson et al. (2020*b*)[61]: CanUD cases: EUR = 17 068; AFR = 3848 ∘ Sherva et al. (2016)[107]: CanUD cases: EUR = 2884; AFR = 1572	∘ Zhou et al. (2020*a*)[127]: OUD: EUR cases = 10 544, EUR opioid-exposed controls = 72 613; AFR cases = 5212, AFR opioid-exposed controls = 26 876	∘ Huggett and Stallings (2020*b*)[55]: CocUD: EUR cases = 2349; AFR cases = 3046 ∘ Gelernter et al. (2014*c*)[34]: provided CocUD case–control data for Huggett and Stallings (2020*b*) analyses

Abbreviations for notable samples incorporating non-European populations: European Ancestry (EUR), African Ancestry (AFR), Latino or Hispanic Ancestry (LAT); East Asian American (EAA); South Asian American (SAA).

Note: numeric superscripts correspond to numbered in-text citations (See Supplementary Material).

[a]At time of review, CocUD sample sizes remain substantially smaller than other SUDs; thus, current CocUD findings and downstream analyses (e.g. h^2_{SNP}, r_g) should be interpreted with caution and require replication in well-powered samples. Efforts to extend CocUD sample sizes are underway.

[b]Denotes findings with SUD or problematic use.

[c]Denotes findings with substance consumption measure.

between alcohol consumption *frequency* and alcohol consumption *quantity* (Mallard et al., 2020; Marees et al., 2020*b*). Specifically, alcohol consumption *quantity* was found to be more genetically similar to AUD and psychopathology, while *frequency* demonstrated negative relationships with AUD and other psychiatric outcomes, and was found to be influenced by measures of SES (Mallard et al., 2020; Marees et al., 2020*b*). Thus, evidence of genetic dissimilarity between consumption and AUD may be being driven by *frequency* of drinking, which in turn, is being influenced by indices of SES. Further studies probing this relationship will be needed to fully disentangle the nuance of the shared and unique genetic etiology across the spectrum of alcohol consumption levels (e.g. normative consumption, binge drinking) and AUD.

The largest genome-wide meta-analysis of copy number variation (CNV) and AUD to date found nine CNV regions (six deletions and three duplications) that were suggestively associated with AUD status in a sample of 3243 cases (Sulovari, Liu, Zhu, & Li, 2018). The most significant association (albeit modest: $p = 3.8 \times 10^{-4}$) was a deletion located on 5q21.3, a region that has previously been reported to be associated in a linkage study examining alcohol craving in a Native American population (Ehlers & Wilhelmsen, 2005).

Nicotine use disorder

Large-scale GWAS of ND have consistently reported GWS associations with cholinergic nicotinic receptor genes. For example, one GWAS reported GWS associations between ND, as assessed by FTND scores, and the well-replicated signal found for genetic variation within the *CHRNA5-CHRNA3-CHRNB4* locus on chromosome 15 (Table 1; Hancock et al., 2018*b*). Hancock et al. (2018*b*) also reported a novel association with an intronic variant (rs910083) in the DNA methyltransferase gene (*DNMT3B*) located on chromosome 20 that was further found to be associated with heavy smoking in the UK Biobank and implicated in the development of lung cancer. A more recently expanded GWAS of ND from the Nicotine Dependence GenOmics Consortium (iNDiGO; Quach et al., 2020) provided further evidence for a top variant association in *CHRNA5* on chromosome 15 (rs16969968), and a significantly associated variant (rs151176846) in *CHRNA4* on chromosome 20 (*CHRNA4*) that was associated with ND in an earlier GWAS (Hancock et al., 2015). The iNDiGO Consortium estimated an h^2_{SNP} of ~0.09 for ND (Quach et al., 2020).

There have also been large-scale efforts examining genetic contributions for other nicotine-related phenotypes (Furberg et al., 2010; Liu et al., 2019). For example, GSCAN (GWAS & Sequencing Consortium of Alcohol and Nicotine use; Liu et al., 2019) reported 467 GWS associations across a variety of smoking-related phenotypes [initiation of regular smoking, quantity of cigarettes per day (CPD), smoking cessation, and age of regular smoking initiation]. In GSCAN, the top single variant association reported for smoking outcomes was between the CPD phenotype and rs16969968 located within *CHRNA5* (Liu et al., 2019), similar to findings reported in other GWAS of smoking behaviors (Furberg et al., 2010) and ND (Hancock et al., 2015, 2018*a*; Quach et al., 2020).

Varying patterns of genetic relationships between NicUD and other smoking phenotypes have been observed. For example, smoking initiation was found to be modestly genetically correlated with ND ($r_g = 0.40$), while CPD and ND were highly genetically correlated ($r_g = 0.95$); this suggests that smoking initiation is less genetically similar to problematic nicotine use relative to CPD (Quach et al., 2020; Table 1). Genetic overlap between CPD (Liu et al., 2019) and ND (Quach et al., 2020) may be being driven, in part, by the fact that iNDiGO ND was assessed using FTND scores, a measure generally accepted as a reasonable assessment for ND but that also includes assessment for the number of cigarettes smoked per day.

There have been efforts examining rare variant associations with nicotine use. A recent exome-chip meta-analysis of 16 studies fine-mapped 124 GWS rare coding variant associations across nicotine use outcomes [i.e. CPD, pack-years (i.e. quantity of cigarette packs smoked in lifetime), smoking initiation, age of smoking initiation; Brazel et al., 2019]. Rare variation accounted for 1.0–2.2% of phenotypic variance across these traits (Brazel et al., 2019).

Cannabis use disorder

There have been fewer replicable genome-wide discoveries for CanUD, due to small sample sizes. To date, the largest GWAS of CanUD ($N_{cases} = 20\,916$), which combined data from iPSYCH, deCODE Genetics, and the PGC (see Box 1), estimated the SNP-heritability (h^2_{SNP}) to be ~0.12 (using an estimated prevalence of 8.5%) and identified two GWS loci: one located on chromosome 7, near the *FOXP2* gene (lead SNP: rs7783012), and the second located on chromosome 8, with brain eQTLs for *CHRNA2* and *EPHX2* (lead SNP: rs4732724; Table 1; Johnson et al., 2020*b*). *FOXP2* plays a role in synaptic plasticity and has been implicated in speech and language development, and the lead risk variant at this locus, rs7783012, has been previously associated with externalizing behaviors. The *CHRNA2* gene, which encodes the α-2 subunit of the neuronal nicotinic acetylcholine receptor, has been previously implicated in GWAS of CanUD (Demontis et al., 2019) as well as tobacco smoking and schizophrenia, both of which are phenotypically and genetically correlated with CanUD (Table 1). The *EPHX2* gene may be involved in the metabolism of cannabinoids, making this an attractive candidate gene for CanUD, but it is currently unclear whether *EPHX2* or *CHRNA2* may be mechanistically responsible for driving the association between this locus and CanUD. Another notable finding is a GWS variant (rs77378271) in the *CSMD1* gene, which has previously been linked to schizophrenia; Sherva et al. identified this variant in their European-ancestries GWAS of DSM-IV cannabis dependence ($N = 8754$) as well as the trans-ancestral meta-analysis ($N = 14\,754$) that included individuals of both European- and African-ancestries (Sherva et al., 2016). However, this gene (as is the case for most CanUD risk variants proposed so far) has not been replicated in other GWAS of CanUD.

Mirroring findings from twin and family studies, GWAS of CanUD have identified significant genetic overlap between CanUD and other SUDs and measures of substance use. CanUD showed significant positive genetic correlations with smoking initiation, ND, CPD, DPW, and AUD (r_g ranging from 0.31 to 0.66; Table 1; Johnson et al., 2020*b*).

Similar to findings for alcohol, recent GWAS of CanUD have found divergence between cannabis use and CanUD, both at the level of individual risk loci as well as genetic relationships with other traits and disorders. Despite a significant correlation of 0.50 between CanUD and lifetime cannabis use, 12 of 22 traits tested had significantly different genetic correlations with CanUD *v.* cannabis use (Johnson et al., 2020*b*). For example,

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

Box 1.

Glossary of relevant terms:

- **Copy number variation (CNV):** A type of structural genetic variation impacting a region of DNA, resulting in the deletion or duplication of genetic information.
- **Endophenotypes:** Any measurable component between genotype and a trait of interest. Useful for examining subfacets and/or transdiagnostic features of SUDs.
- **Epigenome-wide Association Study (EWAS):** Study examining associations between a phenotype and epigenetic markers (e.g. levels of DNA methylation).
- **Expression quantitative trait loci (eQTL):** Genetic loci influencing expression levels of mRNA (messenger RNA) in disease-relevant tissues (e.g. brain eQTLs).
- **Genome-wide Association Study (GWAS):** A genetic study testing associations between a phenotype and genetic variants (measured or imputed) across the genome.
- **Genotype-Tissue Expression (GTEx):** Publicly-available database allowing for the query of a specific genetic variant's involvement in tissue-specific gene expression.
- **Heritability (*h2*)**: The proportion of variation in a phenotype due to genetic factors; traditionally measured using pedigree information (i.e. twin- or family-based studies), but can be assessed using molecular genetic approaches. This is a population-level measure that can vary across time and environments.
- **LD-score regression:** Method that requires only GWAS summary statistics to estimate SNP-based heritability and genetic correlations between traits of interest.
- **Linkage disequilibrium (LD):** The phenomenon wherein nearby genetic variants are inherited non-independently of each other; if two variants are in linkage equilibrium, they are inherited independently of each other (i.e. not correlated or linked).
- **Multi-omics ('omics):** The study of multiple levels of biological information (e.g. epigenome, transcriptome) in a systems-based approach.
- **Pharmacogenomics:** Studies examining how an individual's genetic variation impacts their pharmacological response to medications.
- **Polygenic score (PGS):** An aggregate score of an individual's genetic predisposition for a certain trait, calculated using a 'discovery' GWAS of that phenotype.
- **Rare variation:** Genetic variation occurring at lower frequencies in a population, generally defined as having a minor allele frequency (MAF) <1%.
- **Single nucleotide polymorphism (SNP):** A single base-pair change in the DNA that is relatively common (MAF >1–5%) in the population.
- **SNP-based heritability (*h2*SNP):** The proportion of variation in a phenotype (i.e. heritability) that is accounted for by measured molecular genetic information (i.e. GWAS data). Estimate of the additive genetic variance that can be explained by common SNPs.

Glossary of relevant large-scale genetic efforts for substance use and SUDs:

- **deCODE:** A large, population-based series of studies based out of Reykjavik, Iceland examining genetic influences on complex traits (e.g. SUDs).
- **FinnGen:** A national research initiative leveraging genetic and digital information from Finland national registries to inform biomedicine and personalized healthcare.
- **GSCAN (GWAS & Sequencing Consortium of Alcohol and Nicotine use):** An international genetic consortium aimed toward conducting large-scale meta-analyses of GWAS, low-frequency non-synonymous variation, and whole-genome sequencing studies for alcohol and nicotine use traits.
- **iNDiGO (Nicotine Dependence GenOmics Consortium):** A large-scale genomic consortium examining genetic influences of nicotine dependence and nicotine-related traits across diverse ancestry groups.
- **iPSYCH (Initiative for Integrative Psychiatric Research):** A large national project founded in Denmark in 2012 funded by the Lundbeck Foundation examining genetic and environmental influences for mental health disorders (e.g. SUDs) in over 130 000 Danish individuals.
- **MVP (Million Veteran Program):** One of the world's largest biobanks including genetic, environmental, and medical information from United States Military Veterans.
- **PGC (Psychiatric Genomics Consortium):** A large collaborative effort spanning 800+ investigators, 36 countries, and >400 000 subjects aimed toward elucidating the genetic contributions across psychiatric disorders. The PGC consists of 14 working groups, including the Substance Use Disorders (PGC-SUD) Working Group.
- **UK Biobank:** A large-scale biomedical database and research resource consisting of genetic and health information from greater than 500 000 UK participants.

lifetime cannabis ever-use shows positive genetic correlations with education and age at first birth, and a negative correlation with BMI (+, +, −; Pasman et al., 2018), while CanUD shows genetic correlations in the opposite direction of effect for these three traits (−, −, +; Johnson et al., 2020*b*). This suggests that, while necessary for the development of CanUD, cannabis initiation is at least partly genetically distinct from CanUD.

One recent study used low-pass whole-genome sequencing (WGS) to study CanUD in two samples, one Native American tribal community and one family-based sample of primarily European ancestry (Gizer, Bizon, Gilder, Ehlers, & Wilhelmsen, 2018). Two significant regions were identified in a meta-analysis of the two samples: one protein-coding region, *C1orf110*, and one regulatory region in the *MEF2B* gene.

Opioid use disorder

To date, GWAS of OUD have identified significant loci near the *KCNG2*, *KCNC1*, *APBB2*, *CNIH3*, *RGMA*, and *OPRM1* genes (Table 1; Cheng et al., 2018; Gelernter et al., 2014*b*; Nelson et al., 2016; Polimanti et al., 2020). The largest OUD GWAS to date, conducted in a total of 114 759 individuals (15 756 cases), observed a significant functional coding variant (rs1799971) in the *OPRM1* gene (Zhou et al., 2020*a*, *b*). Other substance use traits (e.g. ever-smoked, alcohol dependence) and psychiatric disorders (e.g. ADHD, schizophrenia) were positively correlated with OUD (Table 1). Zhou et al. (2020*a*, *b*) estimated the h^2_{SNP} of OUD to be 0.11 (S.E. = 0.02).

While there has been less work examining potential differences between the genetic etiology of OUD compared to lifetime ever-use of opioids or non-dependent opioid use, a study from the PGC observed some genetic differences when comparing opioid-dependent individuals, opioid-exposed controls, and opioid-unexposed controls (Polimanti et al., 2020). There were significant relationships between a risk-taking polygenic score (PGS) and the contrast of opioid dependence and unexposed controls, as well as opioid-exposed controls *v.* unexposed controls. A PGS for neuroticism was associated with opioid dependence

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

(contrasted with both unexposed and exposed controls), but not with the unexposed *v.* exposed controls contrast, consistent with the authors' hypothesis that neuroticism is associated with negative affect being related to dependence but not simply exposure (Polimanti et al., 2020). These results suggest that the definition of controls in SUD GWAS needs to be carefully considered.

The largest CNV study of OUD to date identified three common CNVs (two deletions and one duplication) significantly associated with OUD, and several rare CNVs with large effect sizes that reached suggestive levels of evidence (Li et al., 2015). Interestingly, the three significant CNVs (a 18q12.3 deletion, a Xq28 deletion, and a 1q21.3 duplication) were associated with OUD in both the African-ancestries sample ($N_{cases} = 547$) and a combined, trans-ancestral meta-analysis with the European-ancestries samples ($N_{total-cases} = 1601$).

Cocaine use disorder

CocUD GWAS sample sizes have lagged behind those of licit substances. Thus far, there has been one GWS variant identified: rs2629540, located in the *FAM53B* gene (Gelernter et al., 2014*c*). Huggett and Stallings (2020*a*, *b*) applied a gene-wise test to these data and identified four significant genes: *C1QL2*, *STK38*, and *KCTD20* in European Americans ($N = 3176$), and *NDUFB9* in African Americans ($N = 3370$). A meta-analysis of CocUD (all European-ancestry; $N_{cases} = 2085$, $N_{controls} = 4293$) identified an association with *HIST1H2BD* in a gene-based test (Cabana-Domínguez, Shivalikanjli, Fernàndez-Castillo, & Cormand, 2019). They also found positive genetic correlations with schizophrenia, ADHD, major depression, and risk-taking, in line with phenotypic correlations (despite the number of cases being less than recommended for LD score regression; Table 1). Another recent study used cluster analyses to identify CocUD subtypes with reduced phenotypic heterogeneity, one potential barrier to identifying significant genetic variants for psychiatric disorders (Sun, Kranzler, Gelernter, & Bi, 2020). Still, few genetic findings have replicated amongst GWAS of CocUD; we expect the number of robust, replicable findings to increase with larger sample sizes (similar to other SUDs).

The estimated h^2_{SNP} of CocUD is larger than for other SUDs; Huggett and Stallings (2020*a*) estimated $h^2_{SNP} = 0.28$ (S.E. = 0.14) in their genome-wide analysis of CocUD, while Cabana-Dominguez estimated $h^2_{SNP} = 0.27–0.30$ (S.E. = 0.03–0.06), depending on the estimation method (LDSC *v.* GCTA-GREML; Cabana-Domínguez et al., 2019). In contrast, the h^2_{SNP} for the other SUDs discussed here range from 0.07 to 0.12, depending on the phenotype and prevalence (Table 1). The estimations for CocUD heritability may be inaccurate due to under-powered sample sizes – it will be interesting to see if this pattern of larger h^2_{SNP} is borne out in future large-scale GWAS of CocUD.

Clinical and therapeutic implications

Recent efforts have been made to bring molecular genetic findings from large-scale GWAS of SUDs to translational relevance, especially in terms of genetic prediction of SUDs. However, the complexity of SUDs makes genetic prediction efforts difficult and potentially fraught – they are polygenic, heterogeneous, and multifactorial disorders heavily influenced by environmental factors (including access).

PGS have shown promise for the stratification of individuals at risk by their polygenic 'load' for some health conditions; for example, one successful non-SUD application of PGS was reported for coronary disease, where individuals in the highest quintile of genetic risk had an approximately 90% increase in relative risk of experiencing an adverse coronary event compared to individuals in the lowest quintile of genetic risk (Khera et al., 2016). Current SUD PGS explain a relatively small proportion of variance (generally 1–5%) in SUD-related outcomes, especially relative to other known risk factors (SES, SUD family history, comorbid psychiatric disorders; Barr et al., 2020). This limits their current clinical utility. Furthermore, the best-powered GWAS of SUDs to date have been conducted primarily in samples of European ancestries, limiting their predictive utility to individuals who are also of European ancestry (Martin, Daly, Robinson, Hyman, & Neale, 2019). Finally, PGS can be difficult to interpret in layperson's terms (i.e. being in the 95th percentile of polygenic risk for alcohol dependence does not mean you have a 95% chance of developing the disorder). Further research is needed to fully understand the potential benefits, and possible harms, of incorporating genetic information (e.g. PGS) into SUD treatment planning (Driver, Kuo, & Dick, 2020; Lebowitz, 2019; Lebowitz & Ahn, 2018).

Other efforts to utilize molecular genetics for precision medicine purposes have included pharmacogenetic studies, i.e. identifying genetic variability in pharmaceutical treatment response and efficacy. Some GWAS have, post-hoc, identified gene targets for treatment. For example, several genes identified in GWAS of smoking behaviors, including *CHRNA7*, *CHRNA5*, *CHRNA4*, and *CHRNB2*, have been found to moderate the effect of Varenicline, a smoking cessation treatment that operates as a partial agonist at the nicotine acetylcholine a2b4 receptor (King et al., 2012). However, case-only GWAS comparing treatment responders to non-responders are more likely to uncover pharmacogenetic variability than GWAS of the disorder itself. To date, many pharmacogenomic studies of SUDs have been candidate gene-focused (e.g. dopaminergic pathway genes, Patriquin, Bauer, Soares, Graham, & Nielsen, 2015) and have had limited success. However, a recent opioid dosing GWAS identified a variant close to the *OPRM1* locus affecting methadone dosing requirements in African-ancestry individuals (Smith et al., 2017). Another recent GWAS of AUD treatment outcomes identified multiple loci associated with medication-specific outcomes and provided evidence of polygenic contributions to AUD treatment response (Biernacka et al., 2021). The increasing availability of large, longitudinal datasets with access to electronic health records and genotype data may enable more systematic, unbiased investigations into the interactions between genetic variation and medication efficacy (Hartwell & Kranzler, 2019).

Drug repurposing, or identifying a new indication for an existing therapeutic, has emerged as another promising way to bring GWAS findings to therapeutic relevance (So et al., 2017). Bupropion is a classic example of a drug repurposed to some success: while it was originally used as a treatment for depression, clinicians discovered that it aided in smoking cessation (Fava, 2018). In a proof of principle study, So et al. (2017) suggested several repositioning candidates for psychiatric disorders by connecting imputed transcriptomic profiles from GWAS data to drug-induced gene expression profiles, but this has not yet been done for SUDs. While there have not been any success stories to date for repurposed drugs for SUDs discovered using GWAS data, this is an intriguing path forward, particularly for SUDs, where there are still typically few effective pharmaceutical treatments available.

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

At this time, expectations for the use of complex genetics in clinical and therapeutic settings should be tempered; genome-wide data of SUDs are not informative enough to improve upon factors already assessed in the clinic for diagnostic and risk prediction purposes, but there remains the potential for pharmacogenomic and drug repositioning efforts to make an impact on the treatment of SUDs in the future.

Conclusions and future priorities

Molecular genetic studies of SUDs have undergone massive advances during the past 5 years. Increased GWAS sample sizes and the incorporation of additional 'omics data' have contributed to a better understanding of the molecular mechanisms and biological pathways underlying SUDs (Kapoor et al., 2019). In our conclusions, we highlight possible next steps and suggested priorities for the field of SUD genetics:

Increased diversity of SUD GWAS

The majority of GWAS of SUDs to date are composed primarily of individuals of European-ancestry, and thus, the generalizability of these findings to other ancestry groups is uncertain. This gap has the potential to further exacerbate health disparities for individuals of diverse ancestry. This raises the need for efforts to study SUDs in transancestral populations, such as the All of Us Research Program. As shown in Table 1, GWAS of SUDs have included relatively more diverse samples compared to other psychiatric disorders, but the numbers of non-European samples are still well below the European-ancestry sample sizes.

Linkage-disequilibrium patterns differ across populations, which is one reason that discovery GWAS of European ancestry may not lead to maximally-predictive PGS in non-European ancestry target samples (Martin et al., 2019). It is thus imperative, in the interest of scientific discovery and ensuring that everyone benefits equally from those discoveries, that future SUD GWAS focus on increasing the number of samples of non-European ancestry.

Integration of functional genomic data and cross-species translational models

Recent studies have begun to leverage multi-omics data to identify genes and biological processes associated with SUDs. For example, Kapoor et al. (2019) performed differential gene expression analysis on prefrontal cortex tissue from 65 AUD cases and 73 controls, identifying relevant genes and molecular pathways including upregulation of pathways related to immune responses. Markunas et al. (2020) conducted the first epigenome-wide association study (EWAS) of smoking in human post-mortem brain tissue (specifically the nucleus accumbens); they identified seven DNA methylation (DNAm) biomarkers, three of which were located near genes previously implicated as blood-based DNAm biomarkers of smoking and four of which were novel (*ABLIM3*, *APCDD1L*, *MTMR6*, and *CTCF*). Another recent study (Marees et al., 2020*a*) used GTEx (Genotype-Tissue Expression; Ardlie et al., 2015) data to assess the role of eQTLs in six substance use traits; using this approach, they identified novel loci not identified in the original GWAS for five of the traits. Despite demonstrating progress, these studies also highlight current limitations for SUDs, especially the lack of SUD-specific and cell-type-specific multi-omics data sources. For example, while GTEx (Ardlie et al., 2015) is a valuable resource for general tissue-specific gene expression patterns, the data provide no information about substance-induced transcriptomic changes. There is a need for more SUD-specific tissue samples. Evidence from Kapoor et al. (2019) and the Markunas et al. (2020) EWAS further highlight the importance of examining both brain and other tissues (e.g. blood, liver) in substance-specific studies: drugs can have peripheral effects, but brain-specific biomarkers may provide greater insight into the neurobiological effects of substance exposure.

Another potential direction is the integration of human genetic data with findings from animal models of addiction endophenotypes (Reynolds et al., 2020). The substance use genetics literature is rich with rodent models of addictive behaviors (e.g. positive reinforcement via self-administration paradigms, withdrawal avoidance and drug-seeking). Despite the challenges that must be overcome to integrate human and animal genetic data (e.g. handling non-orthologous genes), rodent endophenotypes may provide insight into the neurobiological mechanisms linking genes to SUD risk. Another issue is that there is no certain way to cross-map animal and human phenotypes, limiting the opportunities for translation. However, as a recent proof of principle, one study (Huggett, Bubier, Chesler, & Palmer, 2020*c*) found modest but significant overlap in differentially expressed genes and gene networks when comparing human CocUD case–control data with mice in a cocaine *v.* saline solution self-administration paradigm, suggesting commonalities in the reward circuitry of human CocUD and self-administration paradigms in rodents.

Refinement of phenotypes and ascertainment strategies

A key priority for future genetics studies of SUDs is further examining the implications of broad *v.* deep phenotyping approaches and different sample-ascertainment strategies. Prior studies (reviewed in Sanchez-Roige et al., 2020; Sanchez-Roige & Palmer, 2020) suggest that consumption measures (e.g. alcohol intake frequency, cannabis initiation) have divergent patterns of genetic correlation relative to their respective SUDs. However, contrasts of substance use and use disorder are complicated by several issues, including the recall period: while most measures of SUDs are lifetime diagnoses, measures of substance use are often assessed within a recent timeframe (e.g. past year). Additionally, large, unrepresentative samples (e.g. the UK Biobank is skewed toward older individuals with high SES, the Million Veteran Program is skewed heavily toward males) can lead to collider bias (Munafò, Tilling, Taylor, Evans, & Davey Smith, 2018), biases can arise from misreporting and longitudinal changes (Xue et al., 2021), and there is a complicated interplay between genetic and sociological factors in the context of substance use and the development of SUDs (see larger discussion in previous AUD genetics section; Marees et al., 2020*b*). These issues complicate efforts to examine distinctions between the genetics of substance use and SUDs. As mentioned earlier for alcohol, there are instances in which the genetics of 'use' are intimately linked with the genetics of use disorder (e.g. *ADH1B* and *ALDH2* variants exerting their effects via decreased likelihood of alcohol consumption). Furthermore, there is evidence that both substance use (e.g. lifetime cannabis use, ever smoked cigarettes regularly) and SUDs (e.g. CanUD) are strongly related to general externalizing behaviors (Linnér et al., 2020). Still, a recent preprint identified a common genetic factor that underlies SUDs but is *not* shared with measures of substance use, nor other psychiatric disorders (Hatoum et al., 2021), suggesting that SUDs are not

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

simply the combination of substance exposure and psychopathology. Collectively, these findings suggest the importance of assessing a variety of measures of substance use and SUDs (via clinical diagnoses or shorter questionnaires) from multiple types of samples to further elucidate the genetic architecture of consumption measures compared to problematic use across SUDs.

Another promising direction forward in terms of ascertainment strategy is the development of population-based biobanks with embedded family designs (e.g. FinnGen; https://www.finngen.fi/en). Even in the absence of molecular genetic data, national registries have previously been used in innovative ways to examine the influence of genes and family environment on SUD outcomes in offspring (e.g. the triparental design explored in Kendler, Ohlsson, Sundquist, & Sundquist, 2015). One benefit of genotyped population-based datasets with embedded families is that researchers can apply new genetic methods that leverage relatedness patterns to better understand the ways in which parents may influence children's substance use trajectories both 'directly' (passing on SUD risk alleles) and 'indirectly' (through family environment) (Kong et al., 2018).

Finally, there are multiple substance classes not covered in this review, including hallucinogens, 'club drugs', and inhalants. These substance classes have been included in a handful of twin and family studies examining drug use, but no well-powered GWAS exist. Future GWAS efforts will be informative for how the genetics of these additional SUDs overlap with or diverge from well-studied SUDs.

Conclusion

Recent years have brought substantial progress in advancing our understanding of the genetic architecture of SUDs and other substance use behaviors (e.g. consumption quantity), and relating these findings to etiologically-relevant processes for the development of SUDs. The field will continue to see significant advances in genetic discovery as larger sample sizes of individuals of diverse ancestry begin to become realized. It is the hope that these continued advancements will have clinically meaningful implications for SUD prevention and treatment in the future.

Note

[1] Please see Box 1 for glossary of key terms used in the present review.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291721000969.

Acknowledgements. We thank the co-Principal Investigators of the Psychiatric Genomics Consortium Substance Use Disorders (PGC-SUD) working group, Dr Arpana Agrawal, Dr Howard J. Edenberg, and Dr Joel Gelernter, for their support and for providing critical feedback on this manuscript. We would also like to thank two anonymous reviewers for their thoughtful and constructive comments on an earlier draft of this manuscript.

Financial support. This work was supported by grants from the National Institute of Mental Health under award number MH109532 (J.D.D; E.C.J) and the National Institute on Alcohol Abuse and Alcoholism under award numbers F32AA027435 (E.C.J) and T32AA028259 (J.D.D).

Conflict of interest. None.

References

Abdellaoui, A., Smit, D. J. A., Van Den Brink, W., Denys, D., & Verweij, K. J. H. (2021). Genomic relationships across psychiatric disorders including substance use disorders. *Drug and Alcohol Dependence*, *220*, 108535. doi: 10.1016/j.drugalcdep.2021.108535

Agrawal, A., & Lynskey, M. T. (2006). The genetic epidemiology of cannabis use, abuse and dependence. *Addiction*, *101*(6), 801–812.

Agrawal, A., Neale, M., Jacobson, K., Prescott, C. A., & Kendler, K. S. (2005). Illicit drug use and abuse/dependence: Modeling of two-stage variables using the CCC approach. *Addictive Behavior*, *30*, 1043–1048.

Agrawal, A., Verweij, K. J. H., Gillespie, N. A., Heath, A. C., Lessov-Schlaggar, C. N., Martin, N. G., … Lynskey, M. T. (2012). The genetics of addiction – A translational perspective. *Translational Psychiatry*, *2*(7), e140–e140.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders: Diagnostic and statistical manual of mental disorders, fifth edition*. Arlington, VA: American Psychiatric Association.

Ardlie, K. G., Deluca, D. S., Segre, A. V., Sullivan, T. J., Young, T. R., Gelfand, E. T., … Dermitzakis, E. T. (2015). The genotype-tissue expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science (New York, N.Y.)*, *348*, 648–660.

Barr, P. B., Ksinan, A., Su, J., Johnson, E. C., Meyers, J. L., Wetherill, L., … Dick, D. M. (2020). Using polygenic scores for identifying individuals at increased risk of substance use disorders in clinical and population samples. *Translational Psychiatry*, *10*, 196.

Berrettini, W. (2017). A brief review of the genetics and pharmacogenetics of opioid use disorders. *Dialogues in Clinical Neuroscience*, *19*(3), 229.

Biernacka, J. M., Coombes, B. J., Batzler, A., Geske, J. R., Ho, A. M., Frank, J., … Karpyak, V. M. (2021). Genetic contributions to alcohol use disorder treatment outcomes: A genome-wide pharmacogenomics study. doi:10.1101/2021.02.03.21251107.

Brazel, D. M., Jiang, Y., Hughey, J. M., Turcot, V., Zhan, X., Gong, J., … Surendran, P. (2019). Exome chip meta-analysis fine maps causal variants and elucidates the genetic architecture of rare coding variants in smoking and alcohol use. *Biological Psychiatry*, *85*(11), 946–955.

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., … Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics Nature Genetics*, *47*, 291–295.

Cabana-Domínguez, J., Shivalikanjli, A., Fernàndez-Castillo, N., & Cormand, B. (2019). Genome-wide association meta-analysis of cocaine dependence: Shared genetics with comorbid conditions. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *94*, 109667.

Cheng, Z., Zhou, H., Sherva, R., Farrer, L. A., Kranzler, H. R., & Gelernter, J. (2018). Genome-wide association study identifies a regulatory variant of RGMA associated with opioid dependence in European Americans. *Biological Psychiatry Biological Psychiatry*, *84*, 762–770.

Clarke, T.-K., Adams, M. J., Davies, G., Howard, D. M., Hall, L. S., Padmanabhan, S., … Hayward, C. (2017). Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank ($N = 112\ 117$). *Molecular Psychiatry*, *22*(10), 1376–1384.

Cohen, L. M., Myers, M. G., & Kelly, J. F. (2002). Assessment of nicotine dependence among substance abusing adolescent smokers: A comparison of the DSM-IV criteria and the modified Fagerström tolerance questionnaire. *Journal of Psychopathology and Behavioral Assessment*, *24*, 225–233. doi: 10.1023/A:1020722915204.

Compton, W. M., Dawson, D. A., Goldstein, R. B., & Grant, B. F. (2013). Crosswalk between DSM-IV dependence and DSM-5 substance use disorders for opioids, cannabis, cocaine and alcohol. *Drug and Alcohol Dependence*, *132*(1–2), 387–390.

Deak, J. D., Miller, A. P., & Gizer, I. R. (2019). Genetics of alcohol use disorder: A review. *Current Opinion in Psychology*, *27*, 56–61.

Degenhardt, L., & Hall, W. (2012). Extent of illicit drug use and dependence, and their contribution to the global burden of disease. *The Lancet*, *379* (9810), 55–70.

Demontis, D., Rajagopal, V. M., Thorgeirsson, T. E., Als, T. D., Grove, J., Leppälä, K., … Reginsson, G. W. (2019). Genome-wide association study implicates CHRNA2 in cannabis use disorder. *Nature Neuroscience*, *22*(7), 1066.

Driver, M. N., Kuo, S. I.-C., & Dick, D. M. (2020). Genetic feedback for psychiatric conditions: Where are we now and where are we going. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *183*(7), 423–432.

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

Edenberg, H. J., & Mcclintick, J. N. (2018). Alcohol dehydrogenases, aldehyde dehydrogenases, and alcohol use disorders: A critical review. *Alcoholism Clinical and Experimental Research Alcoholism Clinical and Experimental Research*, *42*, 2281–2297.

Ehlers, C. L., & Wilhelmsen, K. C. (2005). Genomic scan for alcohol craving in Mission Indians. *Psychiatric Genetics*, *15*(1), 71–75.

Fava, M. (2018). The promise and challenges of drug repurposing in psychiatry. *World Psychiatry World Psychiatry*, *17*, 28–29.

Fisher, C. B., & Jaber, R. (2019). Ethical issues in substance-use prevention research. In *Advances in prevention science* (pp. 281–299). Cham, Switzerland: Springer, Cham. doi:10.1007/978-3-030-00627-3_18.

Furberg, H., Kim, Y., Dackor, J., Boerwinkle, E., Franceschini, N., Ardissino, D., … Merlini, P. A. (2010). Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature Genetics*, *42*(5), 441.

Gelernter, J., Kranzler, H. R., Sherva, R., Almasy, L., Koesterer, R., Smith, A. H., … Farrer, L. A. (2014a). Genome-wide association study of alcohol dependence: Significant findings in African- and European-Americans including novel risk loci. *Molecular Psychiatry*, *19*(1), 41–49. doi: 10.1038/mp.2013.145.

Gelernter, J., Kranzler, H. R., Sherva, R., Koesterer, R., Almasy, L., Zhao, H., & Farrer, L. A. (2014b). Genome-wide association study of opioid dependence: Multiple associations mapped to calcium and potassium pathways. *Biological Psychiatry*, *76*(1), 66–74.

Gelernter, J., Sherva, R., Koesterer, R., Almasy, L., Zhao, H., Kranzler, HR., & Farrer, L. (2014c). Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Molecular Psychiatry*, *19*, 717–723.

Gelernter, J., Sun, N., Polimanti, R., Pietrzak, R. H., Levey, D. F., Lu, Q., … Aslan, M. (2019). Genome-wide association study of maximum habitual alcohol intake in >140000 US European and African American veterans yields novel risk loci. *Biological Psychiatry*, *86*(5), 365–376.

Gelernter, J., Zhou, H., Nuñez, Y. Z., Mutirangura, A., Malison, R. T., & Kalayasiri, R. (2018). Genomewide association study of alcohol dependence and related traits in a Thai population. *Alcoholism Clinical and Experimental Research*, *42*(5), 861–868.

Gillespie, N. A., Neale, M. C., & Kendler, K. S. (2009). Pathways to cannabis abuse: A multi-stage model from cannabis availability, cannabis initiation and progression to abuse. *Addiction*, *104*, 430–438. doi: 10.1111/j.1360-0443.2008.02456.x.

Gizer, I. R., Bizon, C., Gilder, D. A., Ehlers, C. L., & Wilhelmsen, K. C. (2018). Whole genome sequence study of cannabis dependence in two independent cohorts. *Addiction Biology*, *23*(1), 461–473.

Grant, B. F., Saha, T. D., Ruan, W. J., Goldstein, R. B., Chou, S. P., Jung, J., … Hasin, D. S. (2016). Epidemiology of DSM-5 drug use disorder: Results from the national epidemiologic survey on alcohol and related conditions-III. *JAMA Psychiatry*, *73*(1), 39–47. doi: 10.1001/jamapsychiatry.2015.2132.

Grotzinger, A. D., Rhemtulla, M., De Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., … Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *3*(5), 513–525. doi: 10.1038/s41562-019-0566-x.

Hancock, D. B., Guo, Y., Reginsson, G. W., Gaddis, N. C., Lutz, S. M., Sherva, R., … Johnson, E. O. (2018a). Genome-wide association study across European and African American ancestries identifies a SNP in DNMT3B contributing to nicotine dependence. *Molecular Psychiatry Molecular Psychiatry*, *23*, 1911–1919.

Hancock, D. B., Markunas, C. A., Bierut, L. J., & Johnson, E. O. (2018b). Human genetics of addiction: New insights and future directions. *Current Psychiatry Reports*, *20*(2), 8.

Hancock, D. B., Reginsson, G. W., Gaddis, N. C., Chen, X., Saccone, N. L., Lutz, S. M., … Zink, F. (2015). Genome-wide meta-analysis reveals common splice site acceptor variant in CHRNA4 associated with nicotine dependence. *Translational Psychiatry*, *5*(10), e651.

Hartwell, E. E., & Kranzler, H. R. (2019). Pharmacogenetics of alcohol use disorder treatments: An update. *Expert Opinion on Drug Metabolism & Toxicology*, *15*(7), 553–564. doi: 10.1080/17425255.2019.1628218.

Hasin, D. S., O'Brien, C. P., Auriacombe, M., Borges, G., Bucholz, K., Budney, A., … Grant, B. F. (2013). DSM-5 criteria for substance use disorders: Recommendations and rationale. *American Journal of Psychiatry*, *170*(8), 834–851. doi: 10.1176/appi.ajp.2013.12060782.

Hatoum, A. S., Johnson, E. C., Polimanti, R., Zhou, H., Walters, R., Gelernter, J., … Agrawal, A. (2021). The addiction genetic factor a (g): A unitary genetic vulnerability characterizes substance use disorders and their associations with common correlates. medRxiv doi:2021.01.26.21250498.

Heath, A. C., Bucholz, K. K., Madden, P. A., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., … Martin, N. G. (1997). Genetic and environmental contributions to alcohol dependence risk in a national twin sample: Consistency of findings in women and men. *Psychological Medicine*, *27*(6), 1381–1396.

Hedegaard, M., Miniño, A., & Warner, M. (2020). NCHS data brief: No. 356: January 2020: Drug overdose deaths in the United States, 1999–2018. Retrieved from https://stacks.cdc.gov/view/cdc/84647-h.pdf.

Huggett, S. B., Bubier, J. A., Chesler, E. J., & Palmer, R. H. C. (2020c). Do meso-limbic gene expression findings from mouse models of cocaine self-administration recapitulate human cocaine use disorder? *Genes, Brain and Behavior*, *20*, e12689. doi: 10.1111/gbb.12689.

Huggett, S. B., & Stallings, M. C. (2020a). Cocaine'omics: Genome-wide and transcriptome-wide analyses provide biological insight into cocaine use and dependence. *Addiction Biology*, *25*(2), e12719.

Huggett, S. B., & Stallings, M. C. (2020b). Genetic architecture and molecular neuropathology of human cocaine addiction. *Journal of Neuroscience*, *40*(27), 5300–5313.

Jang, S.-K., Saunders, G., Liu, M., Jiang, Y., Liu, D. J., Vrieze S., & 23andMe Research Team (2020). Genetic correlation, pleiotropy, and causal associations between substance use and psychiatric disorder. *Psychological Medicine*, 1–11.

Johnson, E. C., Chang, Y., & Agrawal, A. (2020a). An update on the role of common genetic variation underlying substance use disorders. *Current Genetic Medicine Reports*, *8*(2), 35–46.

Johnson, E. C., Demontis, D., Thorgeirsson, T. E., Walters, R. K., Polimanti, R., Hatoum, A. S., … Agrawal, A. (2020b). A large-scale genome-wide association study meta-analysis of cannabis use disorder. *The Lancet Psychiatry*, *7*, 1032–1045.

Jorgenson, E., Thai, K. K., Hoffmann, T. J., Sakoda, L. C., Kvale, M. N., Banda, Y., … Choquet, H. (2017). Genetic contributors to variation in alcohol consumption vary by race/ethnicity in a large multi-ethnic genome-wide association study. *Molecular Psychiatry*, *22*, 1359–1367.

Kapoor, M., Wang, J.-C., Farris, S. P., Liu, Y., Mcclintick, J., Gupta, I., … Goate, A. (2019). Analysis of whole genome-transcriptomic organization in brain to identify genes associated with alcoholism. *Translational Psychiatry*, *9*, 1–11.

Kendler, K. S. (2001). Twin studies of psychiatric illness. *Archives of General Psychiatry Archives of General Psychiatry*, *58*, 1005.

Kendler, K. S., Jacobson, K. C., Prescott, C. A., & Neale, M. C. (2003). Specificity of genetic and environmental risk factors for use and abuse/dependence of cannabis, cocaine, hallucinogens, sedatives, stimulants, and opiates in male twins. *American Journal of Psychiatry*, *160*(4), 687–695.

Kendler, K. S., Karkowski, L. M., Neale, M. C., & Prescott, C. A. (2000). Illicit psychoactive substance use, heavy use, abuse, and dependence in a US population-based sample of male twins. *Archives of General Psychiatry*, *57*(3), 261–269.

Kendler, K. S., Myers, J., & Prescott, C. A. (2007). Specificity of genetic and environmental risk factors for symptoms of cannabis, cocaine, alcohol, caffeine, and nicotine dependence. *Archives of General Psychiatry*, *64*(11), 1313–1320.

Kendler, K. S., Ohlsson, H., Karriker-Jaffe, K. J., Sundquist, J., & Sundquist, K. (2017). Social and economic consequences of alcohol use disorder: A longitudinal cohort and co-relative analysis. *Psychological Medicine*, *47*(5), 925.

Kendler, K. S., Ohlsson, H., Sundquist, J., & Sundquist, K. (2015). Triparental families: A new genetic-epidemiological design applied to drug abuse, alcohol use disorders, and criminal behavior in a Swedish national sample. *American Journal of Psychiatry*, *172*(6), 553–560. doi: 10.1176/appi.ajp.2014.14091127.

Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry*, *60*(9), 929–937.

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

Kessler, R. C. (2004). The epidemiology of dual diagnosis. *Biological Psychiatry*, *56*, 730–737.

Khera, A. V., Emdin, C. A., Drake, I., Natarajan, P., Bick, A. G., Cook, N. R., … Kathiresan, S. (2016). Genetic risk, adherence to a healthy lifestyle, and coronary disease. *New England Journal of Medicine*, *375*(24), 2349–2358. doi: 10.1056/nejmoa1605086.

King, D. P., Paciga, S., Pickering, E., Benowitz, N. L., Bierut, L. J., Conti, D. V., … Park, P. W. (2012). Smoking cessation pharmacogenetics: Analysis of Varenicline and bupropion in placebo-controlled clinical trials. *Neuropsychopharmacology*, *37*, 641–650.

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., … Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science (New York, N.Y.)*, *359*(6374), 424–428. doi: 10.1126/science.aan6877.

Koob, G. F., & Le Moal, M. (2001). Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology*, *24*, 97–129.

Koopmans, J. R., Slutske, W. S., Van Baal, G. C. M., & Boomsma, D. I. (1999). The influence of religion on alcohol use initiation: Evidence for genotype X environment interaction. *Behavior Genetics*, *29*(6), 445–453.

Kranzler, H. R., Zhou, H., Kember, R. L., Smith, R. V., Justice, A. C., Damrauer, S., … Reid, J. (2019). Genome-wide association study of alcohol consumption and use disorder in 274424 individuals from multiple populations. *Nature Communications*, *10*(1), 1–11.

Lebowitz, M. S. (2019). The implications of genetic and other biological explanations for thinking about mental disorders. *Hastings Center Report*, *49*, S82–S87.

Lebowitz, M. S., & Ahn, W. (2018). Blue genes? Understanding and mitigating negative consequences of personalized information about genetic risk for depression. *Journal of Genetic Counseling*, *27*(1), 204–216.

Li, D., Zhao, H., & Gelernter, J. (2012). Strong protective effect of the aldehyde dehydrogenase gene (ALDH2) 504lys (*2) allele against alcoholism and alcohol-induced medical diseases in Asians. *Human Genetics*, *131*, 725–737.

Li, D., Zhao, H., Kranzler, H. R., Li, M. D., Jensen, K. P., Zayats, T., … Gelernter, J. (2015). Genome-wide association study of copy number variations (CNVs) with opioid dependence. *Neuropsychopharmacology*, *40*(4), 1016–1026.

Linnér, R. K., Mallard, T. T., Barr, P. B., Sanchez-Roige, S., Madole, J. W., Driver, M. N., … Dick, D. M. (2020). Multivariate genomic analysis of 1.5 million people identifies genes related to addiction, antisocial behavior, and health. BioRxiv. doi:10.1101/2020.10.16.342501.

Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., … Vrieze, S. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature Genetics*, *51*(2), 237–244.

Mallard, T. T., Savage, J. E., Johnson, E. C., Huang, Y., Edwards, A. C., Hottenga, J. J., … Sanchez-Roige, S. (2020). Multivariate GWAS elucidates the genetic architecture of alcohol consumption and misuse, corrects biases, and reveals novel associations with disease. bioRxiv doi:2020.09.21.304196; doi: https://doi.org/10.1101/2020.09.21.304196.

Marees, A. T., Gamazon, E. R., Gerring, Z., Vorspan, F., Fingal, J., van den Brink, W., … Sherva, R. (2020a). Post-GWAS analysis of six substance use traits improves the identification and functional interpretation of genetic risk loci. *Drug and Alcohol Dependence*, *206*, 107703.

Marees, A. T., Smit, D. J. A., Ong, J.-S., Macgregor, S., An, J., Denys, D., … Derks, E. M. (2020b). Potential influence of socioeconomic status on genetic correlations between alcohol consumption measures and mental health. *Psychological Medicine*, *50*(3), 484–498. doi: 10.1017/s0033291719000357.

Markunas, C. A., Semick, S. A., Quach, B. C., Tao, R., Deep-Soboslay, A., Carnes, M. U., … Hancock, D. B. (2020). Genome-wide DNA methylation differences in nucleus accumbens of smokers vs. nonsmokers. *Neuropsychopharmacology*, *46*, 554–560.

Martin, A. R., Daly, M. J., Robinson, E. B., Hyman, S. E., & Neale, B. M. (2019). Predicting polygenic risk of psychiatric disorders. *Biological Psychiatry*, *86*, 97–109.

Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M., & Davey Smith, G. (2018). Collider scope: When selection bias can substantially influence observed associations. *International Journal of Epidemiology*, *47*(1), 226–235. doi: 10.1093/ije/dyx206.

Nelson, E. C., Agrawal, A., Heath, A. C., Bogdan, R., Sherva, R., Zhang, B., … Montgomery, G. W. (2016). Evidence of CNIH3 involvement in opioid dependence. *Molecular Psychiatry*, *21*, 608–614.

Pasman, J. A., Verweij, K. J., Gerring, Z., Stringer, S., Sanchez-Roige, S., Treur, J. L., … Ong, J.-S. (2018). GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal effect of schizophrenia liability. *Nature Neuroscience*, *21*(9), 1161–1170.

Patriquin, M. A., Bauer, I. E., Soares, J. C., Graham, D. P., & Nielsen, D. A. (2015). Addiction pharmacogenetics: A systematic review of the genetic variation of the dopaminergic system. *Psychiatric Genetics*, *25*(5), 181–193. doi: 10.1097/YPG.0000000000000095.

Payne, T. J., Smith, P. O., McCracken, L. M., McSherry, W. C., & Antony, M. M. (1994). Assessing nicotine dependence: A comparison of the Fagerström tolerance questionnaire (FTQ) with the Fagerström test for nicotine dependence (FTND) in a clinical sample. *Addictive Behaviors*, *19*, 307–317. doi: 10.1016/0306-4603(94)90032-9.

Polimanti, R., Walters, R. K., Johnson, E. C., McClintick, J. N., Adkins, A. E., Adkins, D. E., … Gelernter, J. (2020). Leveraging genome-wide data to investigate differences between opioid use vs. opioid dependence in 41176 individuals from the Psychiatric Genomics Consortium. *Molecular Psychiatry*, *25*, 1673–1687.

Prom-Wormley, E. C., Ebejer, J., Dick, D. M., & Bowers, M. S. (2017). The genetic epidemiology of substance use disorder: A review. *Drug and Alcohol Dependence*, *180*, 241–259. doi: 10.1016/j.drugalcdep.2017.06.040.

Quach, B. C., Bray, M. J., Gaddis, N. C., Liu, M., Palviainen, T., Minica, C. C., … Hancock, D. B. (2020). Expanding the genetic architecture of nicotine dependence and its shared genetics with multiple traits. *Nature Communications*, *11*(1), 1–13. doi:10.1038/s41467-020-19265-z.

Quillen, E. E., Chen, X.-D., Almasy, L., Yang, F., He, H., Li, X., … Gelernter, J. (2014). ALDH2 is associated to alcohol dependence and is the major genetic determinant of 'daily maximum drinks' in a GWAS study of an isolated rural Chinese sample. *American Journal of Medical Genetics Part B Neuropsychiatric Genetics*, *165*, 103–110.

Reynolds, T., Johnson, E. C., Huggett, S. B., Bubier, J. A., Palmer, R. H. C., Agrawal, A., … Chesler, E. J. (2020). Interpretation of psychiatric genome-wide association studies with multispecies heterogeneous functional genomic data integration. *Neuropsychopharmacology*, *46*, 86–97.

Sanchez-Roige, S., Fontanillas, P., Elson, S. L., 23andMe Research Team, Gray, J. C., de Wit, H., … Palmer, A. A. (2019a). Genome-wide association study of alcohol use disorder identification test (AUDIT) scores in 20 328 research participants of European ancestry. *Addiction Biology*, *24*(1), 121–131.

Sanchez-Roige, S., Palmer, A. A., & Clarke, T.-K. (2020). Recent efforts to dissect the genetic basis of alcohol use and abuse. *Biological Psychiatry*, *87*(7), 609–618.

Sanchez-Roige, S., Palmer, A. A., Fontanillas, P., Elson, S. L., 23andMe Research Team, the Substance Use Disorder of the Working Group of the Psychiatric Genomics Consortium, … Clarke, T.-K. (2019b). Genome-wide association study meta-analysis of the Alcohol Use Disorders Identification Test (AUDIT) in two population-based cohorts. *American Journal of Psychiatry*, *176*(2), 107–118.

Sanchez-Roige, S., & Palmer, A. A. (2020). Emerging phenotyping strategies will advance our understanding of psychiatric genetics. *Nature Neuroscience*, *23*, 475–480.

Sherva, R., Wang, Q., Kranzler, H., Zhao, H., Koesterer, R., Herman, A., … Gelernter, J. (2016). Genome-wide association study of Cannabis dependence severity, novel risk variants, and shared genetic risks. *JAMA Psychiatry*, *73*, 472.

Smith, A. H., Jensen, K. P., Li, J., Nunez, Y., Farrer, L. A., Hakonarson, H., … Gelernter, J. (2017). Genome-wide association study of therapeutic opioid dosing identifies a novel locus upstream of OPRM1. *Molecular Psychiatry*, *22*, 346–352.

So, H. C., Chau, C. K., Chiu, W. T., Ho, K. S., Lo, C. P., Yim, S. H., & Sham, P. C. (2017). Analysis of genome-wide association data highlights candidates for drug repositioning in psychiatry. *Nature Neuroscience*, *20*, 1342–1349.

Stone, A. L., Becker, L. G., Huber, A. M., & Catalano, R. F. (2012). Review of risk and protective factors of substance use and problem use in emerging adulthood. *Addictive Behaviors*, *37*(7), 747–775.

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press

Sullivan, P., & Kendler, K. (1999). The genetic epidemiology of smoking. *Nicotine & Tobacco Research*, *1*(1), 51–57. doi: 10.1080/14622299050011811

Sulovari, A., Liu, Z., Zhu, Z., & Li, D. (2018). Genome-wide meta-analysis of copy number variations with alcohol dependence. *The Pharmacogenomics Journal*, *18*(3), 398–405. doi: 10.1038/tpj.2017.35.

Sun, J., Kranzler, H. R., Gelernter, J., & Bi, J. (2020). A genome-wide association study of cocaine use disorder accounting for phenotypic heterogeneity and gene–environment interaction. *Journal of Psychiatry & Neuroscience: JPN*, *45*(1), 34.

Tsuang, M. T., Bar, J. L., Harley, R. M., & Lyons, M. J. (2001). The Harvard twin study of substance abuse: What we have learned. *Harvard Review of Psychiatry*, *9*(6), 267–279.

Tsuang, M. T., Lyons, M. J., Meyer, J. M., Doyle, T., Eisen, S. A., Goldberg, J., … Eaves, L. (1998). Co-occurrence of abuse of different drugs in men. *Archives of General Psychiatry*, *55*, 967.

Van den Bree, M. B., Johnson, E. O., Neale, M. C., & Pickens, R. W. (1998). Genetic and environmental influences on drug use and abuse/dependence in male and female twins. *Drug and Alcohol Dependence*, *52*(3), 231–241.

Verhulst, B., Neale, M. C., & Kendler, K. S. (2015). The heritability of alcohol use disorders: A meta-analysis of twin and adoption studies. *Psychological Medicine*, *45*(5), 1061.

Verweij, K. J. H., Zietsch, B. P., Lynskey, M. T., Medland, S. E., Neale, M. C., Martin, N. G., … Vink, J. M. (2010). Genetic and environmental influences on cannabis use initiation and problematic use: A meta-analysis of twin studies. *Addiction*, *105*, 417–430.

Viken, R. J., Kaprio, J., Koskenvuo, M., & Rose, R. J. (1999). Longitudinal analyses of the determinants of drinking and of drinking to intoxication in adolescent twins. *Behavior Genetics*, *29*(6), 455–461.

Walters, R. K., Polimanti, R., Johnson, E. C., McClintick, J. N., Adams, M. J., Adkins, A. E., … Agrawal, A. (2018). Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nature Neuroscience*, *21*(12), 1656–1669. doi: 10.1038/s41593-018-0275-1.

World Health Organization (2017). *WHO report on the global tobacco epidemic, 2017: Monitoring tobacco use and prevention policies*. Geneva, Switzerland: World Health Organization.

World Health Organization (2018). *Global status report on alcohol and health 2018 Ed*. Geneva, Switzerland: World Health Organization.

Xu, K., Kranzler, H. R., Sherva, R., Sartor, C. E., Almasy, L., Koesterer, R., … Gelernter, J. (2015). Genomewide association study for maximum number of alcoholic drinks in European Americans and African Americans. *Alcoholism, Clinical and Experimental Research*, *39*(7), 1137–1147. doi: 10.1111/acer.12751.

Xue, A., Jiang, L., Zhu, Z., Wray, N. R., Visscher, P. M., Zeng, J., & Yang, J. (2021). Genome-wide analyses of behavioural traits are subject to bias by misreports and longitudinal changes. *Nature Communications*, *12*, 20211. doi: 10.1038/s41467-020-20237-6.

Young, S. E., Rhee, S. H., Stallings, M. C., Corley, R. P., & Hewitt, J. K. (2006). Genetic and environmental vulnerabilities underlying adolescent substance use and problem use: General or specific? *Behavior Genetics*, *36*(4), 603–615. doi: 10.1007/s10519-006-9066-7.

Zhou, H., Rentsch, C. T., Cheng, Z., Kember, R. L., Nunez, Y. Z., Sherva, R. M., … Gelernter, J. (2020a). Association of OPRM1 functional coding variant with opioid use disorder. *JAMA Psychiatry*, *77*(10), 1072–1080.

Zhou, H., Sealock, J. M., Sanchez-Roige, S., Clarke, T.-K., Levey, D. F., Cheng, Z., … Gelernter, J. (2020b). Genome-wide meta-analysis of problematic alcohol use in 435 563 individuals yields insights into biology and relationships with other traits. *Nature Neuroscience*, *23*, 809–818.

https://doi.org/10.1017/S0033291721000969 Published online by Cambridge University Press