Abstract word count: 148
Main word count: 2,783
Introduction section word count: 238
Results word count: 1,201
Discussion word count: 1,344
Methods word count: 3,452
Figures: 2
Tables: 3
Extended Data Figures: 10
Supplemental Tables: 37
References: 39
Methods References: 31

Youth Correlates of Genetic Liability to Substance Use Disorders

Sarah E. Paul,[1] Aaron J. Gorelik,[1] Nicole R. Karcher,[2] Alex P. Miller,[3] David A.A. Baranger,[1,4] Emma C. Johnson,[2] Kimberly H. LeBlanc,[5] Gayathri Dowling,[5] Deanna M. Barch,[1] Alexander S. Hatoum,[1,2] Arpana Agrawal*,[2] Ryan Bogdan*[1]

1. Psychological & Brain Sciences, Washington University in Saint Louis, Saint Louis, Missouri.
2. Department of Psychiatry, Washington University School of Medicine, Saint Louis, Missouri.
3. Department of Psychiatry, Indiana University School of Medicine, Indianapolis, Indiana
4. Department of Psychiatry, Wisconsin College of Medicine, Milwaukee, Wisconsin
5. Division of Extramural Research, National Institute on Drug Abuse, Bethesda, Maryland.

Author's note: Correspondence concerning this article should be addressed to Sarah E. Paul (email: spaul24@wustl.edu, address: CB 1125 Psychological and Brain Sciences Bldg Room 451, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, USA

ABSTRACT

Substance use disorders (SUDs) are moderately-highly heritable and develop in stages that typically begin in adolescence/young adulthood. It remains unclear whether SUD correlates reflect predispositional risk and/or arise as consequences of exposure. Our phenome-wide association study (PheWAS) of polygenic risk for SUDs (i.e., Problematic Alcohol Use [PAU], Tobacco Use Disorder [TUD], Cannabis Use Disorder [CUD], Opioid Use Disorder [OUD], and General Addiction Risk Factor [ADDrf]) in 1,584-5,556 8-13-year-old children enrolled in the Adolescent Brain Cognitive Development (ABCD) Study identified 309 phenotypic associations, particularly among mental health, sleep, screen time, and caffeine consumption phenotypes. Most associations were observed in substance naïve participants. Neuroimaging phenotypes (e.g., resting state functional connectivity) partially accounted for associations between SUD PRS and screen time phenotypes. Genetic liability to SUDs is expressed during childhood as potentially modifiable risk factors that could be targeted to attenuate its expression in adolescence and adulthood.

Main

Substance use disorders (SUDs) are common and heritable (50-60%)[1,2] conditions that significantly adversely impact individuals, families, and society[3–5]. Typically, the first serious symptoms of SUDs arise in late adolescence and/or early adulthood. Identification of the compendium of early markers of genetic susceptibility to SUDs may highlight mechanisms through which predispositional risk manifests that could be targeted to reduce risk.

Here, to identify correlates of genetic liability to SUDs in childhood, we conduct Phenome-wide Association Studies (PheWASs) of polygenic scores (PRS) for Problematic Alcohol Use (PAU), Tobacco Use Disorder (TUD), Cannabis Use Disorder (CUD), Opioid Use Disorder (OUD), and a general Addiction Risk Factor (ADDrf) in individuals genetically similar to African and European reference populations in the Adolescent Brain Cognitive Development Study[SM] (ABCD Study®) at ages 8-11 (baseline; n=1,584, 5,556, respectively) and 10-13 (follow-up; n=1,273, 5,048). We test associations between each PRS and up to 1,697 psychosocial and health-related phenotypes and 421 imaging-derived phenotypes. Follow-up analyses exclude participants who had initiated substance use to rule out potential confounding by concurrent substance use. Ultimately, phenotypes identified as correlates of genetic liability to SUDs during childhood, prior to SUD expression, and in many cases, substance exposures, may reflect putative risk factors through which genetic liability is expressed. These phenotypes could serve as accessible, malleable targets for family-based education/intervention to improve the health of youth at genetic risk for SUDs.

Results

Non-imaging

European-like ancestry. See **Table 1** for sample descriptive information.

Broad Patterns. ***Number of Significant Associations***. Each substance use disorder PRS ($SUD_{PRS;}$ PAU, TUD, CUD, OUD, ADDrf; **Table 2**) was associated with 10-71 baseline (0.79-5.59%) and 19-136 follow-up (1.12-8.01%) phenotypes following Bonferroni correction ($p_{baseline}$<3.93e-05, $p_{follow\text{-}up}$<2.95e-05), with their average absolute standardized betas and ORs ranging from 0.049-0.074 (avg=0.061) and 1.25-1.40 (avg=1.32), respectively (**Figures 1-2; Supplemental Tables 9-10; Extended Data Figures 1-2)**. ***Associations by Domain.*** Across time points but varying by PRS, the screentime/media use, cognition, substance use and perceptions, and school-related domains had the greatest percentage of significantly associated phenotypes (up to 50%; **Table 3**; **Figure 2; Supplemental Note; Supplemental Tables 9-10**).

Child Behavior and Mental Health. SUD_{PRS}--especially CUD_{PRS} and TUD_{PRS}--exhibited multiple associations with mental health phenotypes (0.027≤|B|≤0.069; 1.16≤OR[1]≤1.59), particularly those related to externalizing behavior. Notably, CUD_{PRS} was associated with a significantly greater proportion of child behavior and psychopathology--particularly externalizing--phenotypes compared to all other PRS (baseline $\chi2(1)\geq40.4$, $p\leq3.50e\text{-}08$; follow-up $\chi2(1)\geq23.4$, $p\leq1.31e\text{-}06$). Several non-externalizing associations also emerged, including greater psychotic-like experiences (PAU_{PRS}, TUD_{PRS}, and CUD_{PRS}), youth- or teacher-reported total behavior problems (all except $ADDrf_{PRS}$), and greater sleep problems (e.g., fewer hours, total sleep disturbances, and social jetlag [i.e., discrepancy between weekday and weekend sleep times]).

Screentime/Media Use. Each SUD_{PRS} was positively associated with frequency and time spent using social media. All SUD_{PRS}, except OUD_{PRS}, were associated with weekend screen time and R-rated movie frequency, with additional SUD-specific associations observed (e.g., higher frequency of playing mature video games, as well as weekday screen weekend texting, and weekday TV/movie time (0.038≤|B|≤0.080; 1.15≤OR≤1.26).

Cognition. OUD_{PRS} was negatively associated with more phenotypes related to cognitive performance (e.g., NIH Toolbox total composite, fluid intelligence,) at baseline (n=5) than any other PRS (n≤2) (across waves and PRS, 0.054≤|B|≤0.074). Cognition phenotypes represented 50% of significant baseline phenotypic associations with OUD_{PRS}.

Substance Use and Perceptions. All SUD_{PRS} were associated with caffeine intake (e.g., maximum intake, use summary score, and/or energy drink). Associations also emerged for less perceived harm for regular smokeless tobacco use (all PRS) and more positive cannabis use expectancies (CUD_{PRS}; 0.027≤|B|≤0.083; 1.16≤OR≤1.18).

School-related. Only CUD_{PRS} was associated lower school performance at baseline, and all PRS except $ADDrf_{PRS}$ were linked with lower school performance, increased likelihood of detention, and/or school disengagement, phenotypes at follow-up (0.038≤|B|≤0.088; 1.21≤OR≤1.46).

Prenatal/Postnatal and Early Development. With the exception of OUD_{PRS}, all SUD_{PRS} were associated with unplanned pregnancy and prenatal tobacco exposure; PAU_{PRS} and CUD_{PRS} were additionally associated with prenatal cannabis exposure (0.023≤|B|≤0.035; 1.20≤OR≤1.65).

---

[1] ORs <1 were converted to be ≥1 to facilitate interpretable range descriptions in text.

Family Behavior and Mental Health. All SUD_{PRS} were positively associated with paternal history of drug problems and parental history of depression. Additional positive associations emerged, including with current caregiver externalizing-related and somatic problems (TUD_{PRS} and CUD_{PRS}; $0.025 \leq |B| \leq 0.059$; $1.25 \leq OR \leq 1.46$).

Substance-related Environment. PAU_{PRS}, TUD_{PRS}, CUD_{PRS}, and OUD_{PRS} were associated with the presence of differential cigarette use rules among family members, greater cigarette accessibility, and/or peer tobacco use ($0.029 \leq |B| \leq 0.085$).

Neighborhood/External Environment. At the follow-up wave, CUD_{PRS} was positively associated with the experience of sexual orientation- and weight-related discrimination ($1.36 \leq OR \leq 1.45$); OUD_{PRS} was positively coupled with several neighborhood adversity metrics (e.g., lower high-skill employment and child opportunity; higher number of individuals with disability; $0.029 \leq |B| \leq 0.11$).

Life Events. All PRS were positively associated with stressful life event exposure measured at follow-up; all PRS except OUD_{PRS} were also coupled with greater impact of these events ($0.059 \leq |B| \leq 0.094$; $1.18 \leq OR \leq 1.41$).

Family Environment. SUD_{PRS} were associated with lower parent/caregiver educational attainment (TUD_{PRS}, CUD_{PRS}), a parent/caregiver partner not being a biological parent (PAU_{PRS}, TUD_{PRS}), caregiver-child conflict (TUD_{PRS}, CUD_{PRS}), and less frequent family dinners (PAU; $0.043 \leq |B| \leq 0.067$; $1.18 \leq OR \leq 1.50$).

Peer Environment. At the follow-up wave, PAU_{PRS}, TUD_{PRS}, and CUD_{PRS} were positively associated with involvement with peers who break rules and have been suspended; TUD_{PRS} and CUD_{PRS} were associated with experiencing peer victimization ($OR=1.21$; $0.040 \leq |B| \leq 0.092$).

Physical Health. TUD_{PRS} was negatively associated with the number of extracurricular (largely sports) activities at follow-up ($B=-0.066$).

Physical Development and Identity. No significant associations were observed with any phenotypes within the Physical Development or Identity domains ($B \leq 0.066$, $OR \leq 1.09$, $p \geq 3.72e\text{-}05$, $p_{Bonferroni} \geq 0.063$).

Secondary analyses. ***Substance-naive.*** A majority (76.9%-100%) of Bonferroni-significant associations in the full sample were Bonferroni-significant in the substance-naive subsample (**Supplemental Tables 11-12**). ***SUD_{PRS} Specificity.*** CUD_{PRS}, OUD_{PRS}, and TUD_{PRS}, but not PAU_{PRS} or $ADDrf_{PRS}$, exhibited unique associations when all PRS were entered simultaneously (**Supplemental Note; Supplemental Tables S13-14**). ***Phenotypic Pruning.*** Pruning intercorrelated phenotypes resulted in a domain non-specific monotonic reduction of significant associations (range=0-66.7%; **Supplemental Tables 9-10; Extended Data Figure 3; Supplemental Note**). ***Phenotypic Change.*** No SUD_{PRS}×Time interactions emerged suggesting stability of findings from Baseline to follow-up (**Supplement Note; Supplemental Table 15**). ***Sex Moderation.*** Four SUD_{PRS}×sex interactions were observed (**Supplemental Note; Supplemental Tables 16-17; Extended Data Figure 4**). ***Within-family.*** Few phenotypes were associated with within-family deviations in PRS following Bonferroni correction (i.e., PAU_{PRS}: increased caffeinated soda consumption; TUD_{PRS}: increased caregiver restlessness, greater mood instability; CUD_{PRS}: increased externalizing behavior **Supplemental Tables 18-19; Extended Data Figure 5**). ***Accounting for baseline externalizing***. Over half (58.2%-100%) of follow-up phenotypes that were previously Bonferroni-significant remained associated with SUD_{PRS} after covarying for

baseline CBCL externalizing scores (**Supplemental Table 20**). These notably included measures related to caffeine intake, screentime, and sleep disturbance.

African-like ancestry. See **Table 1** for sample descriptive information.
Seven follow-up (and no baseline) phenotypes, primarily externalizing-related behaviors, were associated with SUD_{PRS} (**Figure 1c, d; Extended Data Figures 6-7; Supplemental Tables 21-22**). All associations were significantly correlated with associations observed in EUR-like ancestry for AUD/PAU_{PRS}, TUD_{PRS}, CUD_{PRS}, and $ADDrf_{PRS}$ (r=0.089-0.39, p≤1.45e-04). Significant associations remained significant in the substance-naive subsample (**Supplemental Tables 23-24**). Secondary analyses revealed: 1) unique associations between CUD_{PRS} and externalizing behaviors (**Supplemental Tables 25-26**), 2) no significant SUD_{PRS}×Time or SUD_{PRS}×Sex interactions (**Supplemental Tables 27-29**), and e) that all CUD_{PRS} associations remained significant in within-family analyses (**Supplemental Note; Supplemental Tables 30-31**).

Brain Imaging Results. Results were only significant in the EUR-like subsample. SUD_{PRS} were linked to RSFC (greater sensorimotor hand and sensorimotor mouth network connectivity [PAU_{PRS}, CUD_{PRS}]; frontoparietal and sensorimotor hand networks [CUD_{PRS}]) and DWI (lower fractional anisotropy in the right corticospinal/pyramidal [TUD_{PRS}], superior corticostriate [TUD_{PRS}], and left superior corticostriate-parietal cortex [TUD_{PRS}] tracts) metrics. These associations remained in the substance-naive subsample (**Supplement; Supplemental Tables 32-33**. ***Indirect effects results.*** All RSFC and DWI metrics significantly, indirectly linked SUD_{PRS} to follow-up phenotypes (**Supplemental Note; Extended Data Figures 8-10; Supplemental Tables 34-35**).

Cross-Ancestry Meta-Analysis. Between 5.71 and 64.3% of Bonferroni-significant associations in EUR were similarly significant in the cross-ancestry meta-analysis. Many associations that did not survive correction in the meta-analysis exhibited heterogeneous effects across ancestry groups (**Supplemental Note; Supplemental Tables 36-37**).

Discussion

Genetic liability to substance use disorders (i.e., PAU_{PRS}, TUD_{PRS}, CUD_{PRS}, OUD_{PRS}, $ADDrf_{PRS}$) was strongly and most frequently associated with elevated externalizing behavior in both the EUR- and AFR-like individuals, and with screentime/media use, caffeine intake, sleep problems, family history of substance use and depression, and prenatal factors (e.g., unplanned pregnancy, prenatal substance exposure) among EUR in our PheWAS of 8-13-year-old children. Most associations persisted among substance-naive youth and thus are likely to reflect early markers of genetic liability to SUD, rather than—or in addition to—causal consequences of substance use. Collectively, this characterization of environmental, behavioral, and neural correlates of genetic liability to SUDs during childhood prior to onset of substance involvement highlights putative mechanisms through which genetic liability to substance involvement may manifest (see[6,7] for related work in adults). Certain potentially modifiable correlates, such as screentime/media, caffeine use, and sleep should be investigated as candidates for preventative policy and intervention seeking to attenuate SUD risk.

SUD_{PRS} were strongly associated with childhood externalizing (e.g., conduct problems, aggression, inattention). Indeed, the few significant associations in AFR almost entirely comprised externalizing and related (e.g., impulsivity, school performance) phenotypes, several of which also exhibited significant cross-ancestry consistency unlike other phenotypes that were specific to EUR. These externalizing associations may overlapping genetic architecture[8] and heterotypic continuity wherein early externalizing behaviors (e.g., ADHD, oppositional defiant disorder symptoms) represent an earlier manifestation of broad externalizing predisposition, which later includes SUDs[9].

Caffeine intake, screentime/media use, and sleep problems—which were particularly robustly associated with SUD_{PRS}—have been previously genetically and/or phenotypically correlated with externalizing problems[10–15]. However, our externalizing-adjusted PheWAS revealed that SUD_{PRS} associations with follow-up caffeine intake, screentime, and sleep disturbance were partially independent. Whether these robust residual associations are due to pleiotropy that is exclusive to SUD genetic liability, predisposition toward emotion dysregulation distinct from externalizing, and/or to other explanations is worth exploring, particularly in light of recent work in ABCD documenting prospective associations between problematic screen use and substance initiation[16]. For instance, gene-environment correlation between genetic liability to SUDs and a lack of parental supervision could contribute to excessive engagement with media and mature content viewership. In addition, youth at heightened genetic liability to SUDs were less likely to report reading for pleasure. Future studies could thus examine the extent to which the deleterious associations of media engagement could be mitigated by participation in prosocial (e.g., sporting activities, hobby clubs) and scholastic (e.g., reading, art) engagement[17].

Reflecting the ubiquity of gene-environment correlation, genetic liability to SUDs was associated with prenatal factors (e.g., tobacco exposure, unplanned pregnancy), having rule breaking peers, experiencing victimization from and aggressing upon peers as well as experiencing stressful life events, living in a disadvantaged neighborhood and having increased ease of access to cigarettes. Particularly during childhood, genetic liability is intertwined with passive exposure to environments shaped by those genetic susceptibilities, exacerbating their

dual impact[18]. These early, risky, and genotype-related environments have been linked to later substance use[19] and variability in brain development[20] that might hinder the emergence of the cognitive restraint necessary to deter progression to SUD, making them ideal targets for public health interventions. These associations also amplify the need to account for gene-environment correlation in research on genetic and environmental factors. To some degree, our follow-up analyses within pairs of twins and siblings evaluated this direct contribution of genetic susceptibility beyond environments experienced similarly by siblings. We found no evidence for within-family associations, although the reduced sample size combined with constrained PRS variability may have reduced statistical power. Nonetheless, given the predominance of familial environment in the etiology of initial stages of substance use[21], many of our associations may reflect indirect pathways, from genotype to behavior, via shaping of the childhood environment.

SUD_{PRS} was robustly linked to sleep problems and less nightly sleep. Beyond the well-established acute effects of psychoactive substances and withdrawal on sleep[22,23], growing evidence highlights a shared genetic liability underlying both sleep disruptions and substance-related phenotypes[24], including traits like evening chronotype[25(p201)], consistent with our social jetlag findings[26]. Whether these findings reflect yet another aspect of behavioral dysregulation, are consequent to the anti-somnolent effects of caffeine intake and protracted exposure to screens, and/or are attributable to loci with pleiotropic effects on SUD and sleep, our results suggest that youth with genetic susceptibility to SUDs should be monitored for sleep difficulties and their resultant impact on health.

Consistent with small, difficult-to-detect brain-behavior associations[27], few statistically significant associations between SUD_{PRS} and imaging metrics emerged. However, there were two notable findings. *First*, both PAU_{PRS} and CUD_{PRS} were associated with increased resting-state functional connectivity (RSFC) between the sensorimotor hand and sensorimotor mouth networks, and CUD_{PRS} was additionally linked with increased RSFC between the frontoparietal and sensorimotor hand networks. These neural metrics indirectly linked these SUD_{PRS} to screen time behavior. A recent meta-analysis found some evidence that increased sensorimotor cortex connectivity is associated with SUD diagnosis but also with less craving, better SUD treatment response, and reduced consumption among those with SUD[28], and another study found positive associations with the presence of gaming disorder[29], suggesting a potential shared pathway for multiple "addictive" behaviors. *Second*, TUD_{PRS} was associated with reduced white matter integrity in the superior cortical-striatal, corticospinal/pyramidal, and cortical-striate-parietal tracts, tracts connecting regions frequently implicated in theoretical and clinical studies of addiction[30,31], and these brain metrics partially mediated PRS association with screentime behavior. As findings were evident in substance-naive youth, they raise the intriguing possibility of preexisting brain variability associated with SUD genetic susceptibility[31].

Lastly, while there is compelling support for genetic commonality across SUDs, captured here by ADDrf, some SUD-specific distinctions are noteworthy. CUD_{PRS} was associated with significantly more mental health phenotypes, which may be due to its high loading with other externalizing constructs[8,32] or due to the nature of electronic health record diagnoses, which for CUD, likely reflect complicating additional psychopathology[33]. On the other hand, OUD_{PRS} was associated with fewer phenotypes overall, possibly due to its smaller size, but

was uniquely associated with lower global cognitive performance and greater environmental adversity, converging with prior genetic studies[34].

There are some limitations to the study. *First*, despite our inclusion of data from individuals of African-like ancestry, sample sizes of these discovery GWAS and the target AFR-like ABCD sample were small. Summary statistics for other global ancestral groups was absent. *Second*, data from Million Veterans Program, which over-represents older male veterans, are a major component of the SUD GWAS from which PRS were derived, potentially influencing associations; however, these GWAS appear to genetically correlate with GWAS from other samples (e.g., rg=0.66-1.00[35,36]), and our analyses of sex differences did not yield significant results. *Third*, there was little endorsement of substance use beyond alcohol sipping[37], so we were unable to examine associations with substance use phenotypes beyond alcohol sipping and caffeine consumption. It is also possible that early substance experimentation is not strongly related to SUD genetic liability, and/or that it is better explained by genetic factors related to externalizing and sensation seeking[38,39]. *Fourth*, we elected to use Bonferroni correction, which given substantial inter-phenotype correlations, may be overly conservative. Because of this potential for false negatives, FDR-corrected and uncorrected p-values are also presented. *Fifth*, it is worth noting that interrogation of specific mechanisms that link SUD_{PRS} to individual behaviors (e.g., SUD_{PRS} influencing peer rule-breaking via gene-environment correlation) were beyond the scope of the current study, and alternative mechanisms to those proposed herein are also likely.

Conclusions

By identifying associations between genetic liability to SUDs and conditions observable well before the onset of problematic (or any) substance use, the current PheWAS provides insights into childhood behaviors and environments that co-aggregate with and may be harbingers of future problems. While the current analyses do not test causal hypotheses regarding the role of these childhood phenotypes on later SUD, nor can PRS alone provide complete prognostic insight into multifactorial phenotypes like SUD, many of the associated phenotypes, such as caffeine consumption, screen time and social media use, sleep problems, and cigarette availability represent modifiable targets, interventions for which would broadly benefit youth health, and likely even more so in genetically susceptible youth.

Acknowledgments
SEP, AA, and RB conceptualized the study. SEP, AJG, NRK, and APM curated and verified the data for analysis. SEP, DAAB, and ASH developed the R code. SEP conducted analyses and generated tables and figures. SEP, AA, and RB drafted the manuscript. SEP, AJG, NRK, APM, DAAB, ECJ, KHL, GD, DMB, ASH, AA, and RB critically reviewed the manuscript and take responsibility for the decision to submit the manuscript for publication.

Code availability:
Code is available upon request from the corresponding author.

Data Availability
Data used in the preparation of this Article were obtained from the Adolescent Brain Cognitive Development (ABCD) study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9–10 years and follow them over 10 years into early adulthood. The ABCD data repository grows and changes over time.

Funding
This work was funded by the National Institute on Drug Abuse of the National Institutes of Health under grant number R01DA054750 (RB, AA). Data for this study were provided by the Adolescent Brain Cognitive Development (ABCD) study, which was funded by awards U01DA041022, U01DA041025, U01DA041028, U01DA041048, U01DA041089, U01DA041093, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, and U24DA041147 from the NIH and additional federal partners (https://abcdstudy.org/federal-partners.html). Authors received additional funding support from NIH and NSF: F31AA029934 (SEP), K23MH12179201 (NRK), DGE-213989 (AJG), K01DA051759 (ECJ), K01AA030083 (ASH), K99AA030808 (DAAB). R01-MH113883 (DMB), R01-MH066031 (DMB), U01-MH109589 (DMB), U01-A005020803 (DMB), R01-MH090786 (DMB), U01DA055367 (RB), R01DA046224 (RB), R01 HD113188 (RB). The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Dr. Gayathri Dowling was substantially involved in all the cited grants. Dr. Kimberly LeBlanc was substantially involved in U24DA041147 and U24DA041123 consistent with her role as a Scientific Officer. The views and opinions expressed in this manuscript are those of the authors only and do not necessarily represent the views, official policy or position of the U.S. Department of Health and Human Services or any of its affiliated institutions or agencies. This manuscript is the result of funding in whole or in part by the National Institutes of Health (NIH). It is subject to the NIH Public Access Policy. Through acceptance of this federal funding, NIH has been given a right to make this manuscript publicly available in PubMed Central upon the Official Date of Publication, as defined by NIH.

Conflicts of Interests
The authors declare no conflicts of interest.

Inclusion & Ethics:

Due to a dearth of GWAS in populations other than those genetically most similar to European and African reference populations, the present analyses were restricted to these individuals. The results may not generalize to other populations. In addition, the GWAS in African-like ancestry are smaller and thus less well-powered, likely limiting the discovery of phenotypes associated with SUD genetic risk in the African-like ancestry subsample.

The ABCD Study is a large sample of youth, which is an understudied population.

Table 1. Sample descriptive statistics.

Variable	African ancestry (n=1,584)	European ancestry (n=5,556)
	Mean (SD) / n (%)	**Mean (SD) / n (%)**
Sex (female)	792 (50.0%)	2,612 (47.0%)
Age (years)	9.90 (0.60)	9.93 (0.63)
Household Income		
< $35,000	725 (54.2%)	375 (7.1%)
$35,000 - $49,999	176 (13.2%)	283 (5.3%)
$50,000 - $74,999	192 (14.3%)	717 (13.5%)
$75,000 - $99,999	97 (7.2%)	896 (16.9%)
$100,000 - $199,999	124 (9.3%)	2,178 (41.1%)
≥ $200,000	24 (1.8%)	849 (16.0%)
Highest Caregiver Education		
Less than high school	141 (8.9%)	25 (0.45%)
High school degree/equivalent	400 (25.3%)	188 (3.4%)
Some college/associate degree	621 (39.3%)	1046 (18.8%)
College degree	214 (13.6%)	1,753 (31.6%)
Master's degree	167 (10.6%)	1,723(31.0%)
Doctorate/professional degree	36 (2.3%)	820 (14.8%)

Note. Sample descriptive information from the baseline wave of the Adolescent Brain Cognitive Development (ABCD) Study.

Table 2. Discovery GWAS Used to Generate PRS

GWAS Phenotype	Ancestry and n
Alcohol Use Disorder(AUD)/Problematic Alcohol Use (PAU[3a]	AUD: African n=122,571 supplemented with European n=739,895 PAU: European n=903,147
Tobacco Use Disorder (TUD[35])	African n=114,420[b] European n=739,895
Cannabis Use Disorder (CUD[36])	African n=123,208[b]; European n=886,025
Opioid Use Disorder (OUD[34])	African n=84,877[b]; European n=554,186
Addiction Risk Factor (ADDrf[40])	African n=92,630[b]; European n=1,025,550

Note. [a]The cross-ancestry genetic-effect correlation between PAU and AUD is 0.71[3]. [b]African ancestry samples were supplemented with the European ancestry sample in PRS-CSx.

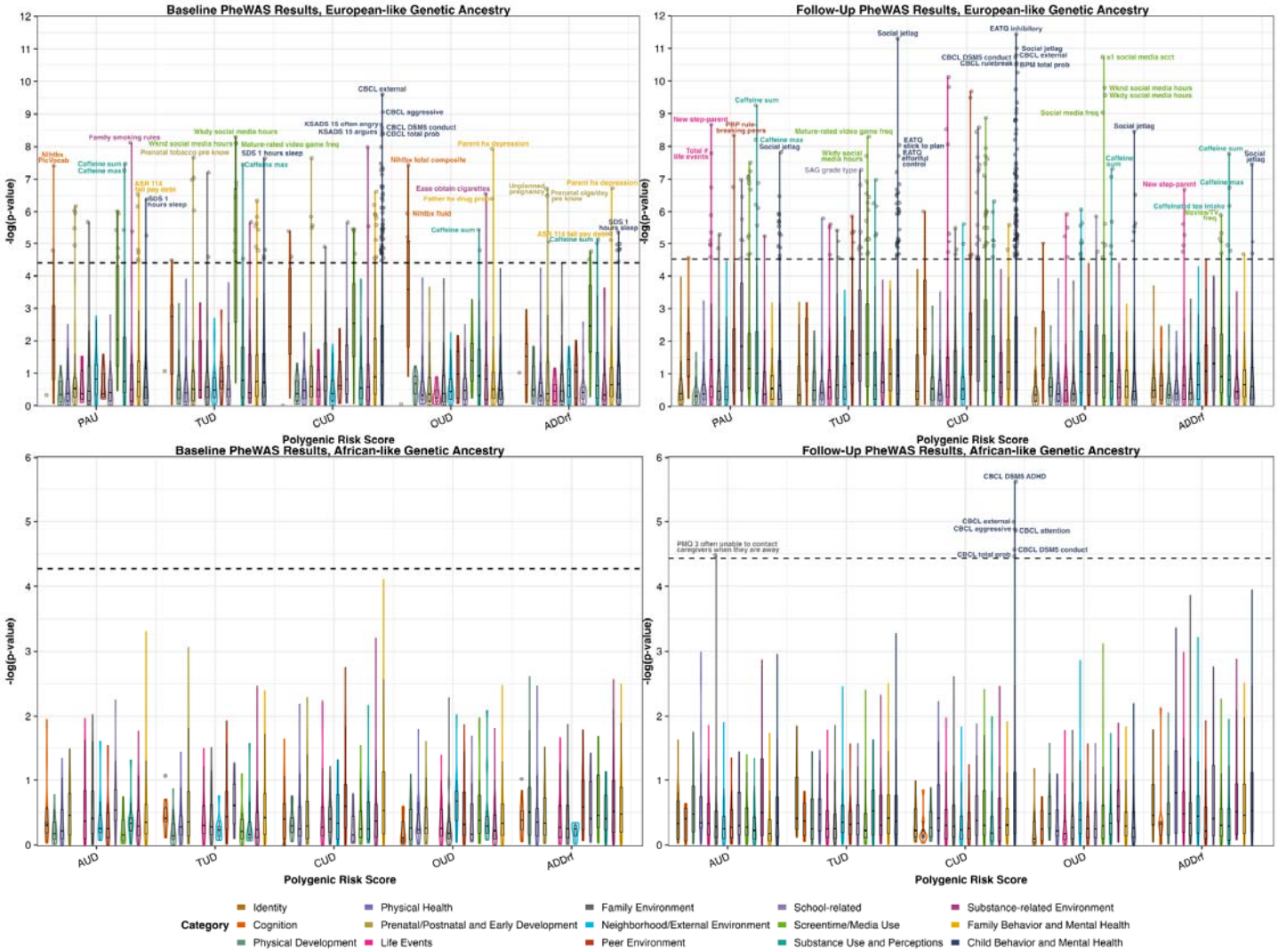
Table 3. Summary of associations by phenotype domain.

		Baseline									
		Problematic Alcohol Use PRS		Tobacco Use Disorder PRS		Cannabis Use Disorder PRS		Opioid use Disorder PRS		Addiction Risk Factor PRS	
	Total	FDR n(%)	Bonf n(%)	FDR n(%)	Bonf n(%)	FDR n(%)	Bonf n(%)	FDR n(%)	Bonf n(%)	FDR n(%)	Bonf n(%)
Child Behav. & MH	634	37 (5.8)	2 (0.32)	90 (14.2)	4 (0.63)	239 (37.7)	48 (7.6)	23 (3.6)	0 (0.0)	38 (6.0)	7 (1.1)
Fam Behav. & MH	239	20 (8.4)	6 (2.5)	40 (16.7)	7 (2.9)	70 (29.3)	9 (3.8)	12 (5.0)	2 (0.84)	12 (5.0)	3 (1.3)
Sub-related Environ.	24	4 (16.7)	2 (8.3)	4 (16.7)	1 (4.2)	5 (20.8)	2 (8.3)	3 (12.5)	1 (4.2)	4 (16.7)	0 (0.0)
Sub Use & Perceptions	29	5 (17.2)	3 (10.3)	7 (24.1)	2 (6.9)	4 (13.8)	0 (0.0)	4 (13.8)	2 (6.9)	6 (20.7)	2 (6.9)
Screentime & Media Use	18	8 (44.4)	5 (27.8)	15 (83.3)	9 (50.0)	11 (61.1)	4 (22.2)	2 (11.1)	0 (0.0)	9 (50.0)	2 (11.1)
School-related	34	1 (2.9)	0 (0.0)	6 17.6)	0 (0.0)	9 (26.5)	1 (2.9)	0 (0.0)	0 (0.0)	1 (2.9)	0 (0.0)
Peer Environ.	11	0 (0.0)	0 (0.0)	1 (9.9)	0 (0.0)	2 (18.2)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Neighborhood/ Ext Environ.	48	2 (4.2)	0 (0.0)	3 (6.3)	0 (0.0)	1 (2.1)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Fam Environ.	74	3 (4.1)	1 (1.4)	8 (10.8)	2 (2.7)	19 (25.7)	1 (1.4)	4 (5.4)	0 (0.0)	5 (6.8)	0 (0.0)
Life Events	6	0 (0.0)	0 (0.0)	2 (33.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Prenat./Postnat. & Early Dev.	60	7 (11.7)	5 (8.3)	7 (11.7)	4 (6.7)	9 (15.0)	4 (6.7)	4 (6.7)	0 (0.0)	7 (11.7)	3 (5.0)
Physical Health	61	1 (1.6)	0 (0.0)	4 (6.6)	0 (0.0)	4 (6.6)	0 (0.0)	2 (3.3)	0 (0.0)	2 (3.3)	0 (0.0)
Physical Dev.	19	0 (0.0)	0 (0.0)	2 (10.5)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Cognition	13	4 (30.8)	2 (15.4)	8 (61.5)	1 (7.7)	9 (69.2)	2 (15.4)	8 (61.5)	5 (38.5)	2 (15.4)	0 (0.0)
Identity	1	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
		Two-Year Follow-Up									

		Problematic Alcohol Use PRS		Tobacco Use Disorder PRS		Cannabis Use Disorder PRS		Opioid use Disorder PRS		Addiction Risk Factor PRS	
	Total	FDR n(%)	Bonf n(%)	FDR n(%)	Bonf n(%)	FDR n(%)	Bonf n(%)	FDR n(%)	Bonf n(%)	FDR n(%)	Bonf n(%)
Child Behav. & MH	663	104 (15.7)	13 (2.0)	172 (25.9)	31 (4.7)	261 (39.4)	81 (12.2)	42 (6.3)	6 (0.9)	48 (7.2)	3 (0.5)
Fam Behav. & MH	158	9 (5.7)	0 (0.0)	31 (19.6)	0 (0.0)	45 (28.5)	2 (1.3)	6 (3.8)	0 (0.0)	6 (3.8)	1 (0.63)
Sub-related Environ.	52	6 (11.5)	1 (1.9)	7 (13.5)	0 (0.0)	9 (17.3)	0 (0.0)	4 (7.7)	0 (0.0)	3 (5.8)	0 (0.0)
Sub Use & Perceptions	75	21 (28.0)	6 (8.0)	17 (22.7)	5 (6.7)	24 (32.0)	6 (8.8)	13 (17.3)	5 (6.7)	11 (14.7)	5 (6.7)
Screentime & Media Use	107	30 (28.0)	11 (10.3)	51 (47.7)	13 (12.1)	44 (41.1)	16 (15.0)	26 (24.3)	8 (7.5)	25 (23.4)	5 (4.7)
School-related	45	19 (42.2)	6 (13.3)	19 (42.2)	7 (15.6)	27 (60.0)	13 (28.9)	14 (31.1)	2 (4.4)	12 (26.7)	0 (0.0)
Peer Environ.	38	12 (31.6)	2 (5.3)	14 (36.8)	5 (13.2)	19 (50.0)	5 (13.2)	7 (18.4)	0 (0.0)	7 (18.4)	0 (0.0)
Neighborhood/ Ext Environ.	123	7 (5.7)	0 (0.0)	15 (12.2)	0 (0.0)	11 (8.9)	2 (1.6)	31 (25.2)	9 (7.3)	8 (6.5)	0 (0.0)
Fam Environ.	111	9 (8.1)	2 (1.8)	18 (16.2)	3 (2.7)	23 (20.7)	3 (2.7)	7 (6.3)	0 (0.0)	6 (5.4)	0 (0.0)
Life Events	162	25 (15.4)	9 (5.6)	16 (9.9)	3 (1.9)	26 (16.0)	7 (4.3)	11 (6.8)	4 (2.5)	16 (9.9)	5 (3.1)
Physical Health	84	1 (1.2)	0 (0.0)	6 (7.1)	1 (1.2)	4 (4.8)	0 (0.0)	2 (2.4)	0 (0.0)	0 (0.0)	0 (0.0)
Physical Dev.	31	0 (0.0)	0 (0.0)	1 (3.2)	0 (0.0)	2 (6.5)	0 (0.0)	1 (3.2)	0 (0.0)	1 (3.2)	0 (0.0)
Cognition	8	2 (25.0)	1 (12.5)	4 (50.0)	0 (0.0)	5 (62.5)	1 (12.5)	3 (37.5)	1 (12.5)	1 (12.5)	0 (0.0)
Identity	40	4 (10.0)	0 (0.0)	4 (10.0)	0 (0.0)	9 (22.5)	0 (0.0)	1 (2.5)	0 (0.0)	2 (5.0)	0 (0.0)

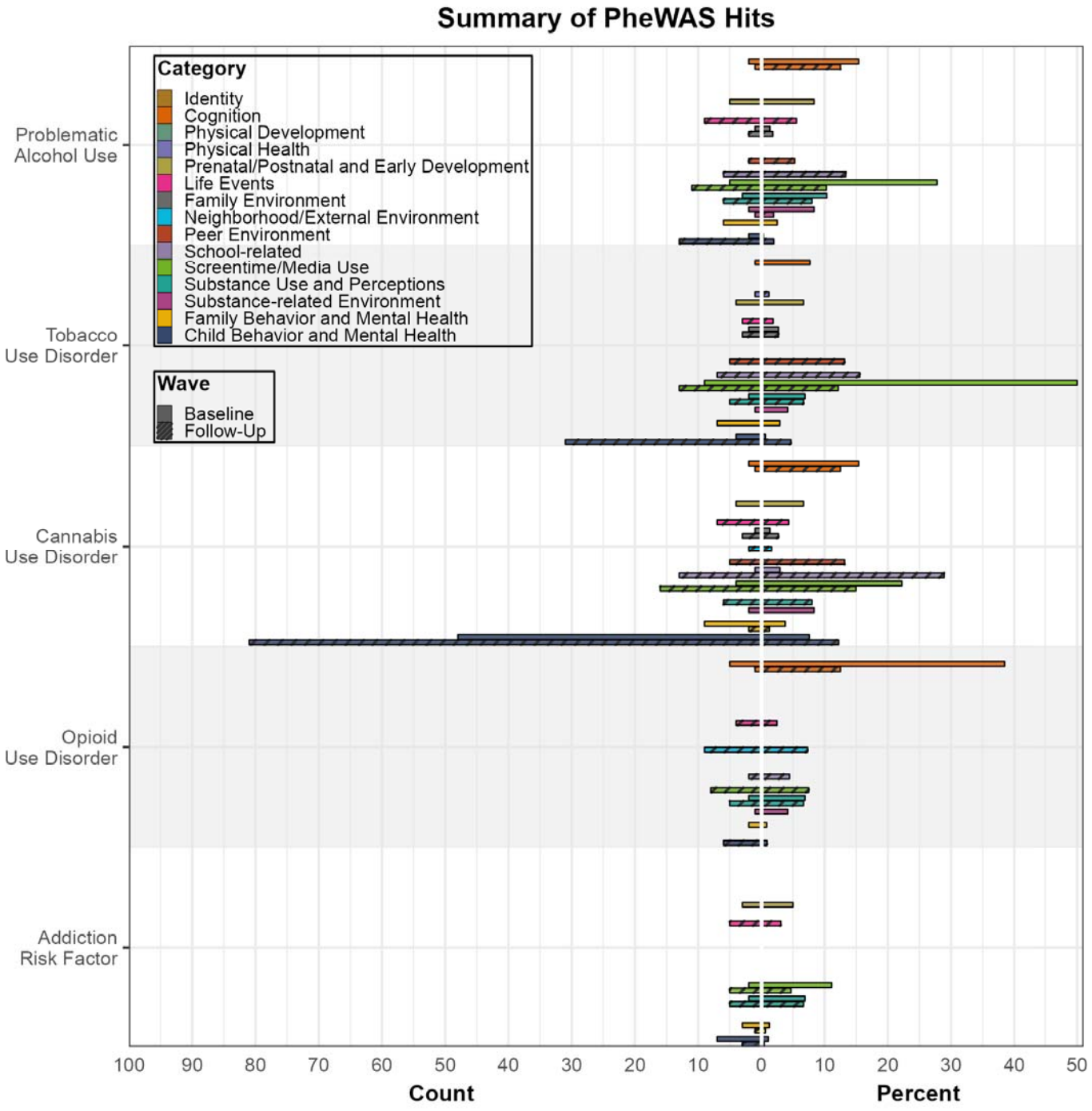
Note. Summary of the number of significant associations per domain, polygenic risk score (PRS), and assessment timepoint. Percents in parentheses are the percent of significantly associated phenotypes of the total number of phenotypes assessed in that domain. Behav = Behavior; MH = Mental Health; Fam = Family; Sub = Substance; Environ = Environment; Ext = External; Dev = Development; FDR = False Discovery Rate; Bonf = Bonferroni.

Figure 1.



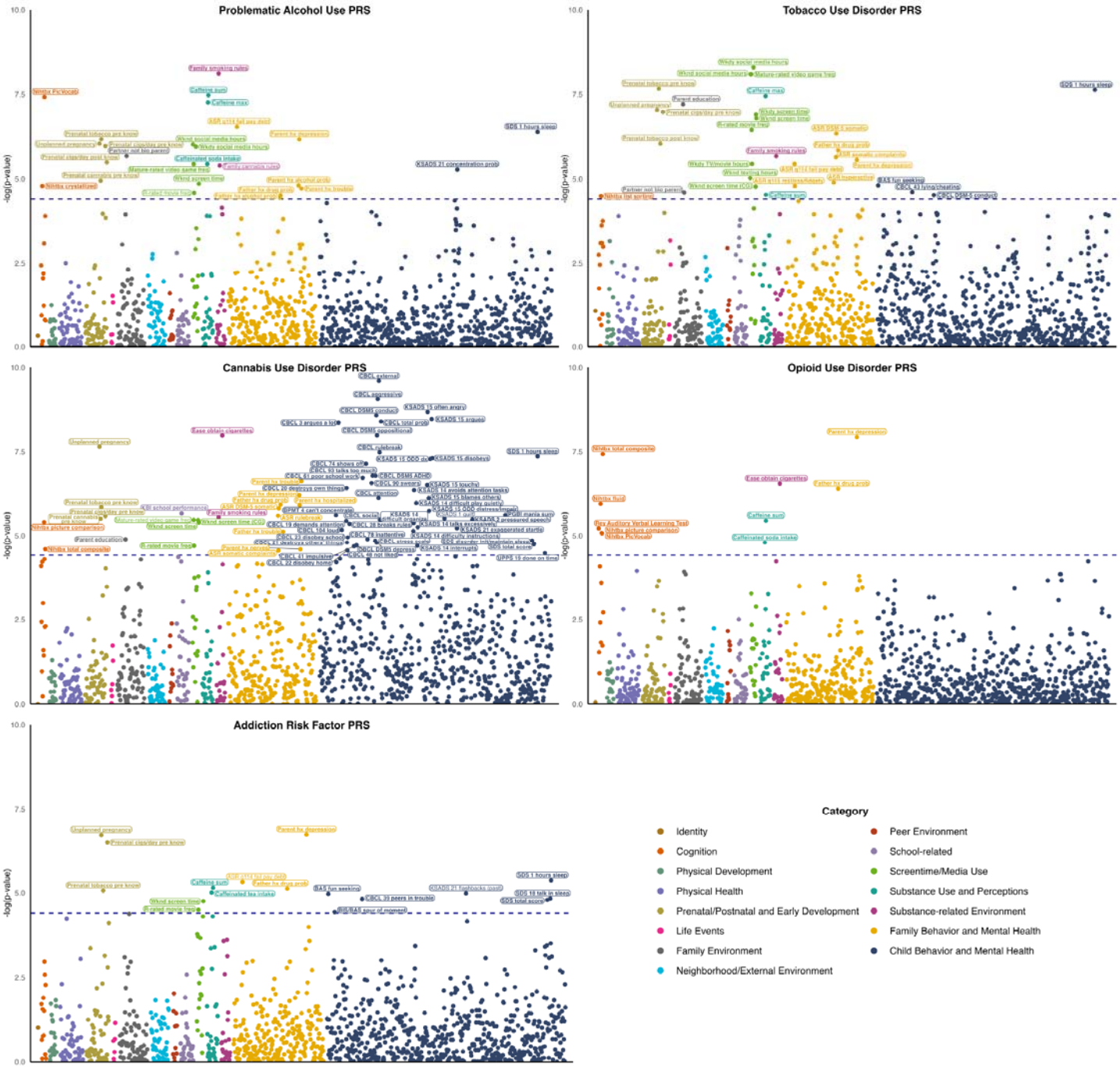
Note. Manhattan and violin plots depicting the number, distribution, and types of phenotypes, within each category, associated with each of the 5 examined PRS. The x-axes are the 5 PRS, the y-axes are the negative log of the p-value of association (higher value indicates stronger association), and color indicates phenotypic Category. The dashed horizontal lines denote the Bonferroni-significant threshold (i.e., 0.05/number of phenotypes). The violin plots represent the distribution of -log(p-values) within each Category, separately for each PRS. Boxplots within each violin plot show the median -log(p-value of association), and the boxes encompass the middle 50% of associations. PAU = Problematic Alcohol Use; TUD = Tobacco Use Disorder; CUD = Cannabis Use Disorder; OUD = Opioid Use Disorder; ADDrf = Addiction Risk Factor; AUD = Alcohol Use Disorder.

Figure 2.



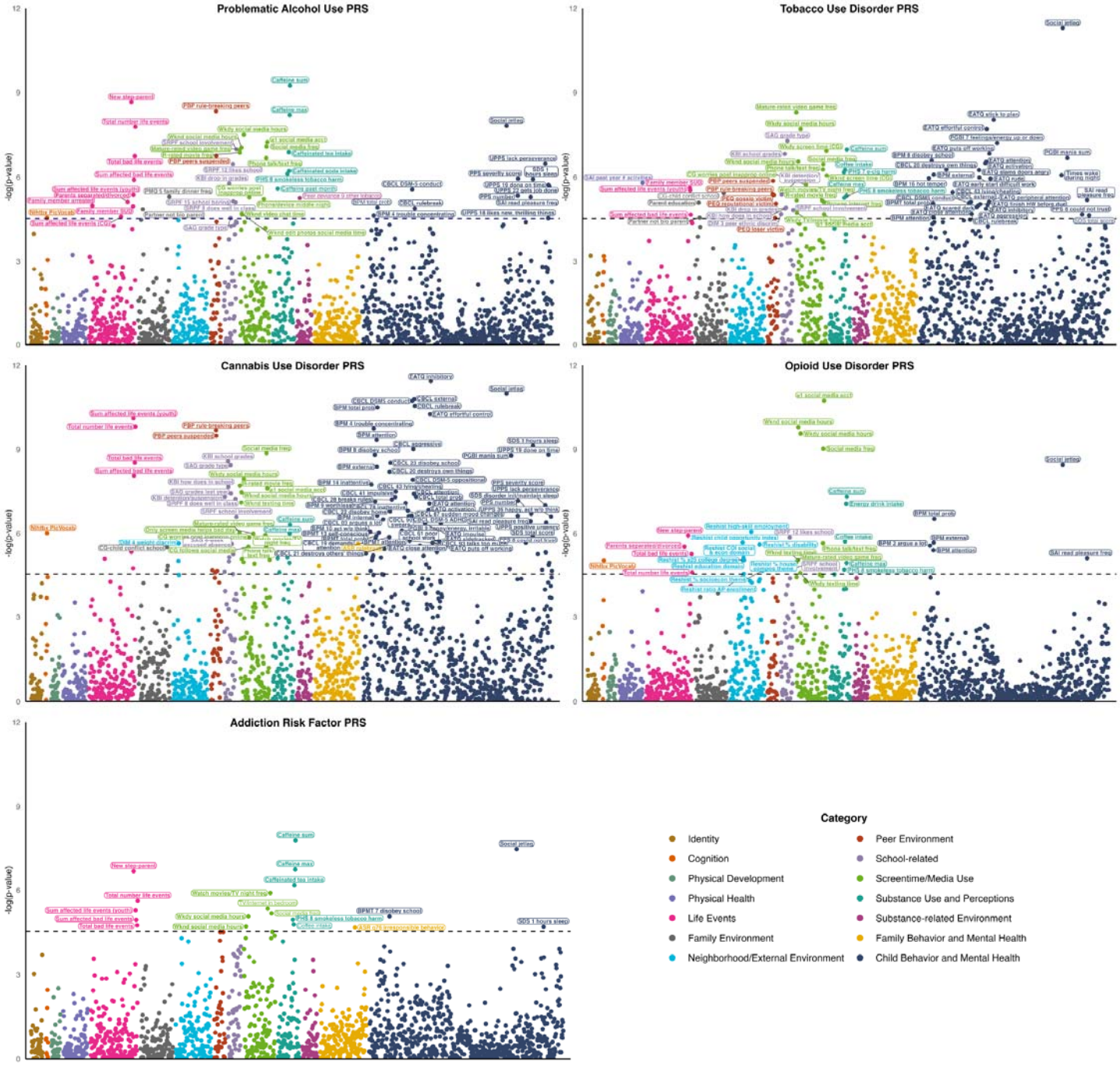
Note. Summary of the number ("Count," left) and percent (right) of associations observed between each PRS (y-axis) and phenotypes in each category, for European-like genetic ancestry. Colors denotes different categories, and solid and dashed bars represent Baseline and Two-Year Follow-Up associations, respectively. Percent reflects the percent of phenotypes in a category associated with a particular PRS of the total number of phenotypes in that category.

Extended Data Figure 1.



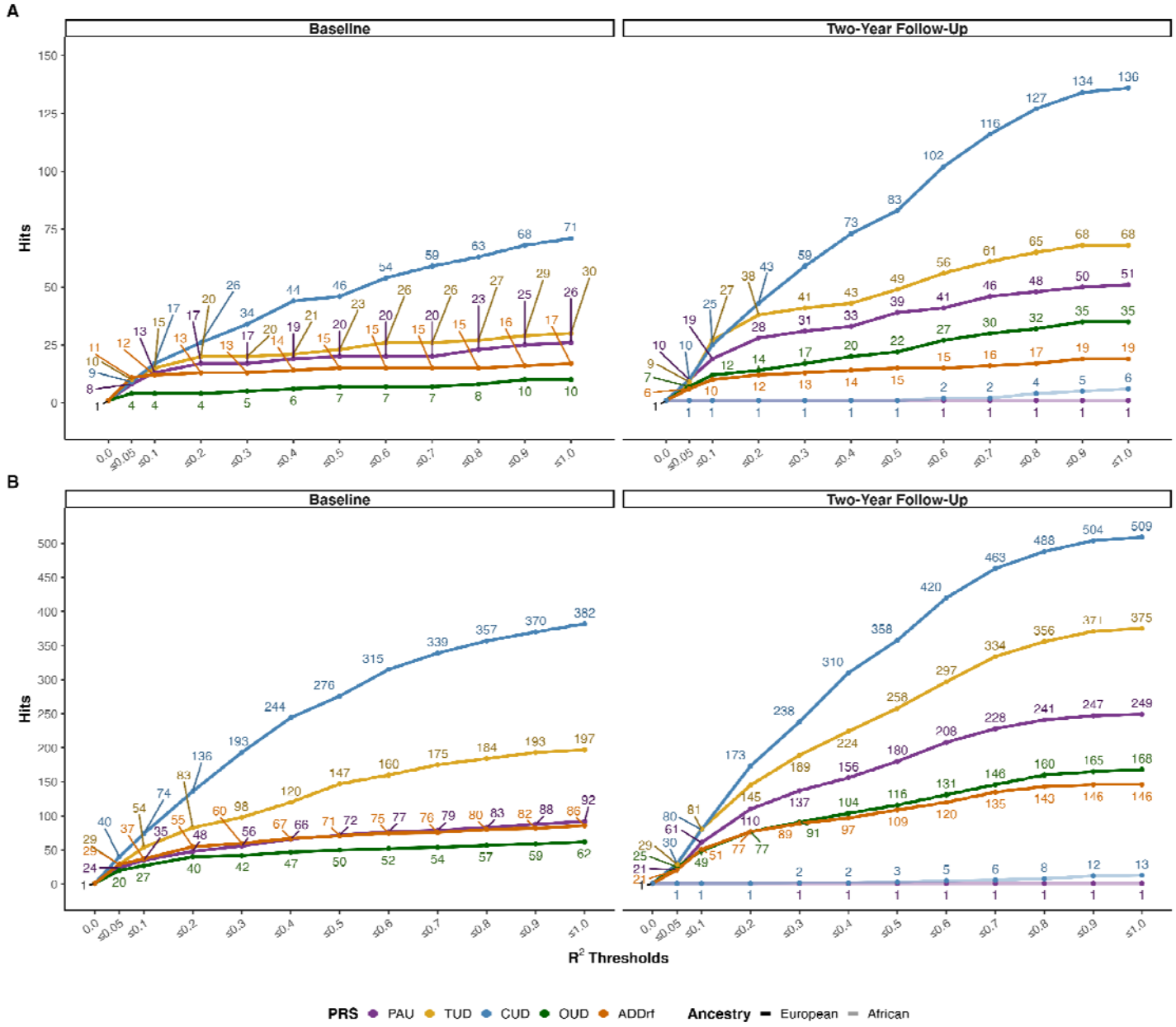
Note. Manhattan plots depicting primary PheWAS associations at the baseline wave among individuals genetically similar to European reference populations. The 1,271 examined phenotypes are spread along the x-axis, and the y-axes represent the -$\log_{10}$ of the uncorrected *p* values of associations between each PRS and each phenotype. The points are color-coded by phenotype domain. The dashed lines indicate the Bonferroni-significant threshold (i.e., -$\log_{10}(0.05/1271) = 4.405$). Full results are located in **Supplemental Table 9**.

Extended Data Figure 2.



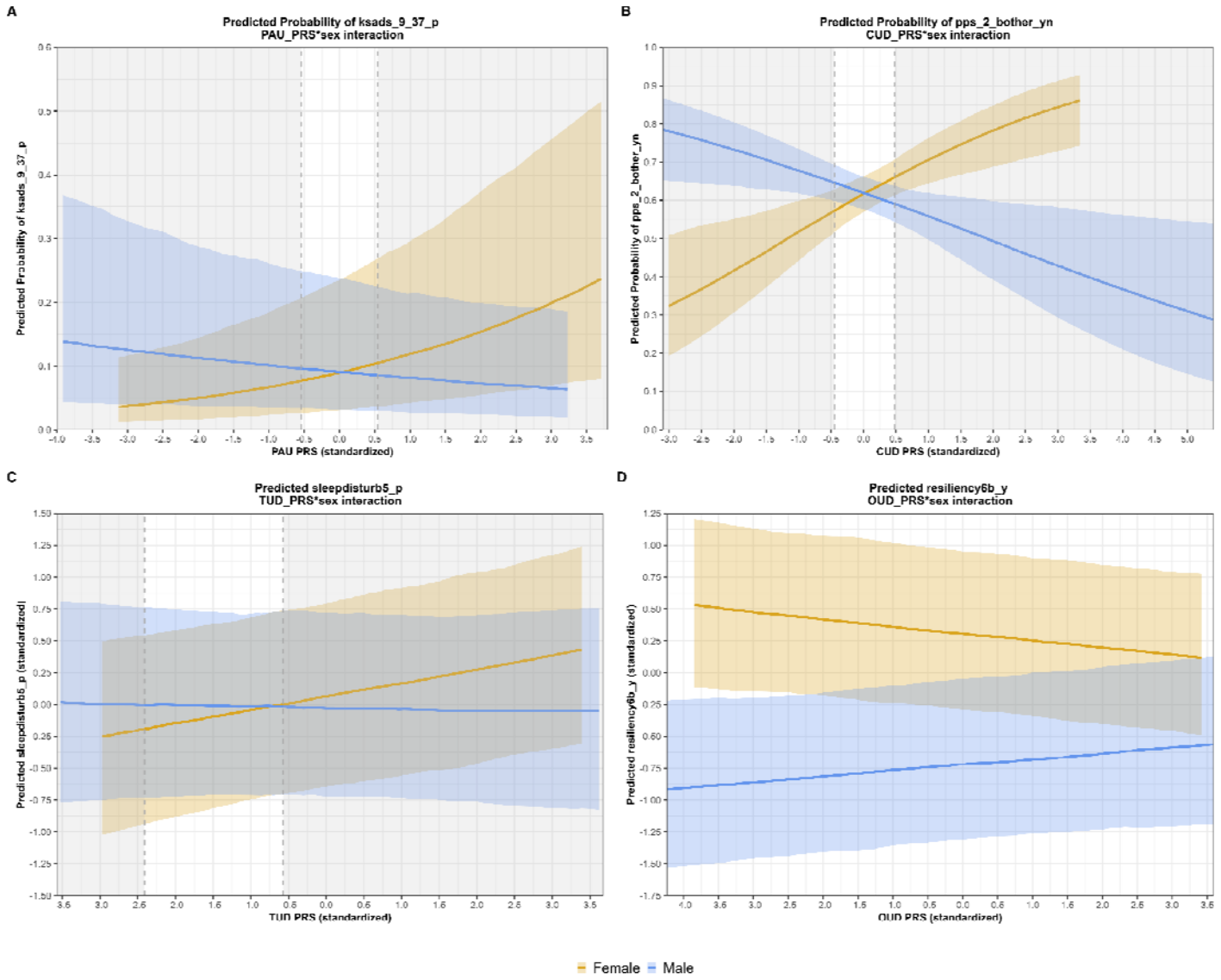
Note. Manhattan plots depicting primary PheWAS associations at the two-year follow-up wave among individuals genetically similar to European reference populations. The 1,697 examined phenotypes are spread along the x-axis, and the y-axes represent the -$\log_{10}$ of the uncorrected *p* values of associations between each PRS and each phenotype. The points are color-coded by phenotype domain. The dashed lines indicate the Bonferroni-significant threshold (i.e., -$\log_{10}(0.05/1697) = 4.53$). Full results are located in **Supplemental Table 10**.

Extended Data Figure 3.



Pruning significant associations according to correlations among phenotypes and strength of association between each phenotype and each PRS. Pseudo R^2 values were obtained for each pair of phenotypes significantly (**A**, Bonferroni-corrected; **B**, FDR-corrected) associated with each PRS. At each R^2 threshold, the phenotype with the higher (i.e., less significant) p-value within each phenotype pair whose R^2 was above the threshold was removed. The following R^2 thresholds were examined: 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, with 1.0 including all significantly associated phenotypes. Colors denote different PRS. More transparent lines denote associations in individuals most genetically similar to African reference populations; these are only included in the Two-Year Follow-Up plots given the lack of significant associations observed at Baseline. PAU = Problematic Alcohol Use; TUD = Tobacco Use Disorder; CUD = Cannabis Use Disorder; OUD = Opioid Use Disorder; ADDrf = Addiction Risk Factor; AUD = Alcohol Use Disorder.

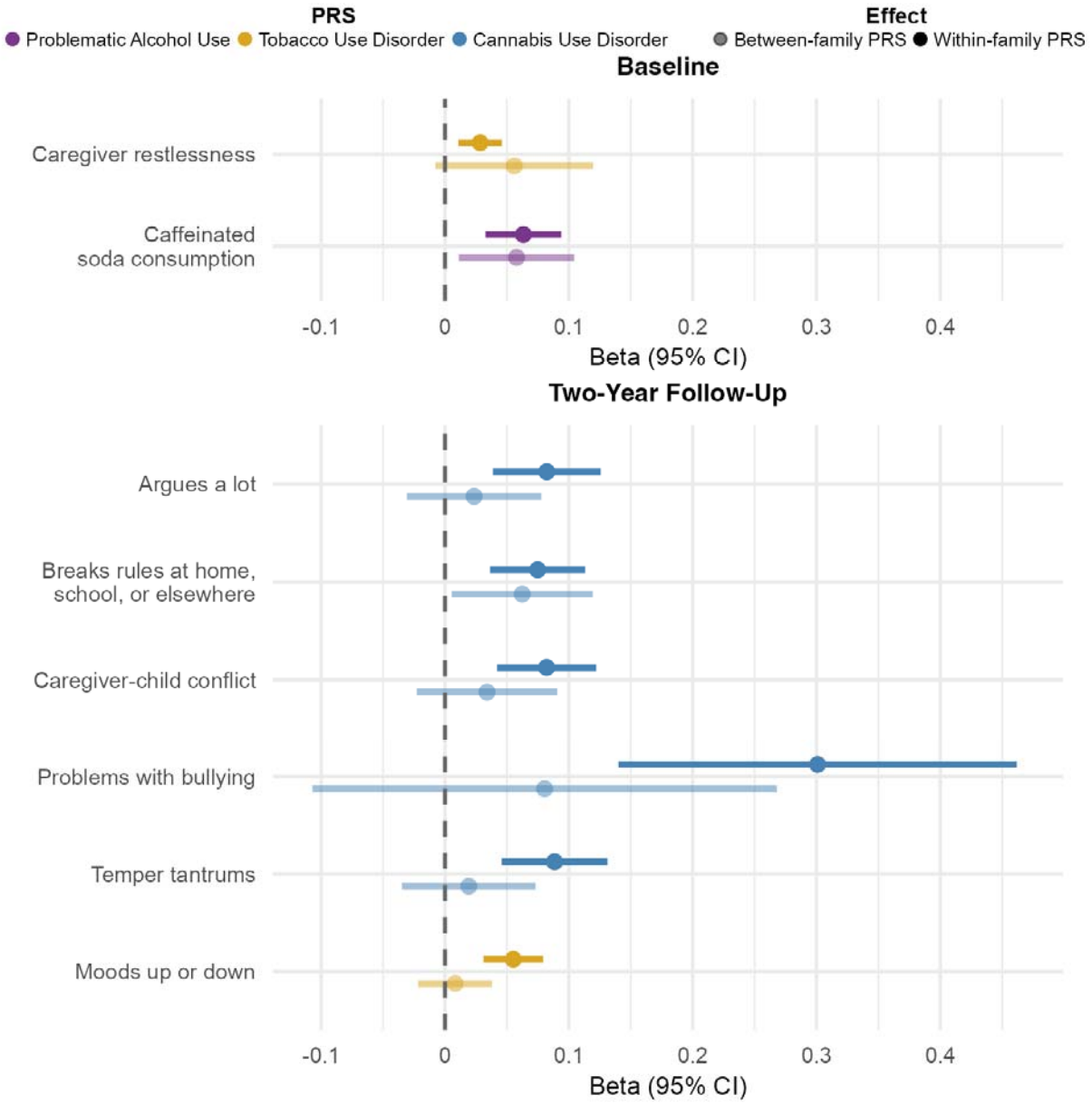
Extended Data Figure 4.



Note. Interactions between SUD PRS and biological sex in the sample of participants genetically similar to European reference populations. Confidence intervals around the predicted probabilities are shown in yellow (female) and blue (male). Gray shaded areas indicate Johnson-Neyman intervals for which the interaction is significant. **A.** Interaction between Problematic Alcohol Use (PAU) PRS and biological sex in association with ksads_9_37_p (past active avoidance of a phobic object). Confidence intervals around the predicted probabilities are shown in yellow (female) and blue (male). The simple slope for female participants is statistically significant (B=0.31, *p*=5.12e-05), and the simple slope for male participants is not (B=-0.12, *p*=0.09). The interaction is significant outside of the range -0.54 to 0.54. **B.** Interaction between Cannabis Use Disorder (CUD) PRS and biological sex in association with pps_2_bother_yn (bothered by hearing strange sounds). The simple slopes for female (B=0.42, *p*=2.16e-06) and male (B=-0.24, *p*=0.02) participants are both statistically significant but opposite in direction. The interaction is significant outside of the range of -0.45 to 0.48. **C.** Interaction between Tobacco Use Disorder (TUD) PRS and biological sex in association with sleepdisturb5_p (feeling anxious or disturbed when falling asleep). The simple slope for female participants is statistically significant (B=0.10, *p*=5.63e-06), and the simple slope for male participants is not (B=-0.02, *p*=0.41). The interaction is significant outside of the range of -2.41 to -0.57. **D.** Interaction between Opioid Use Disorder (CUD) PRS and biological sex in association with resiliency6b_y (number of close friends that are girls). The simple slope for male participants is statistically significant (B=0.08, *p*=3.22e-07), and the simple slope for

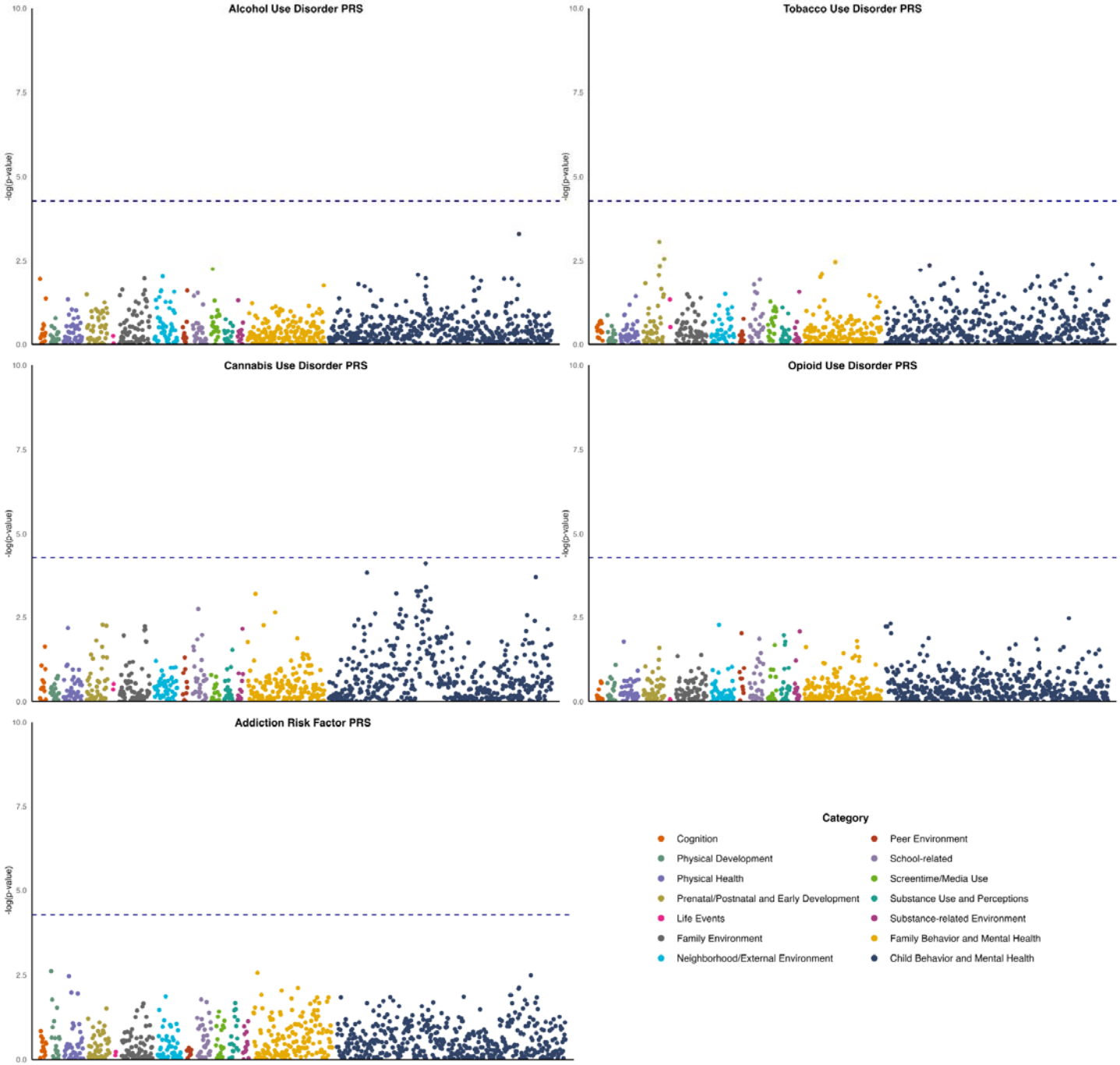
female participants is not (B=-0.02, *p*=0.26). The Johnson-Neyman Intervals for which the interaction is significant fall outside of the range of observed data and thus are not shown here.

Extended Data Figure 5.



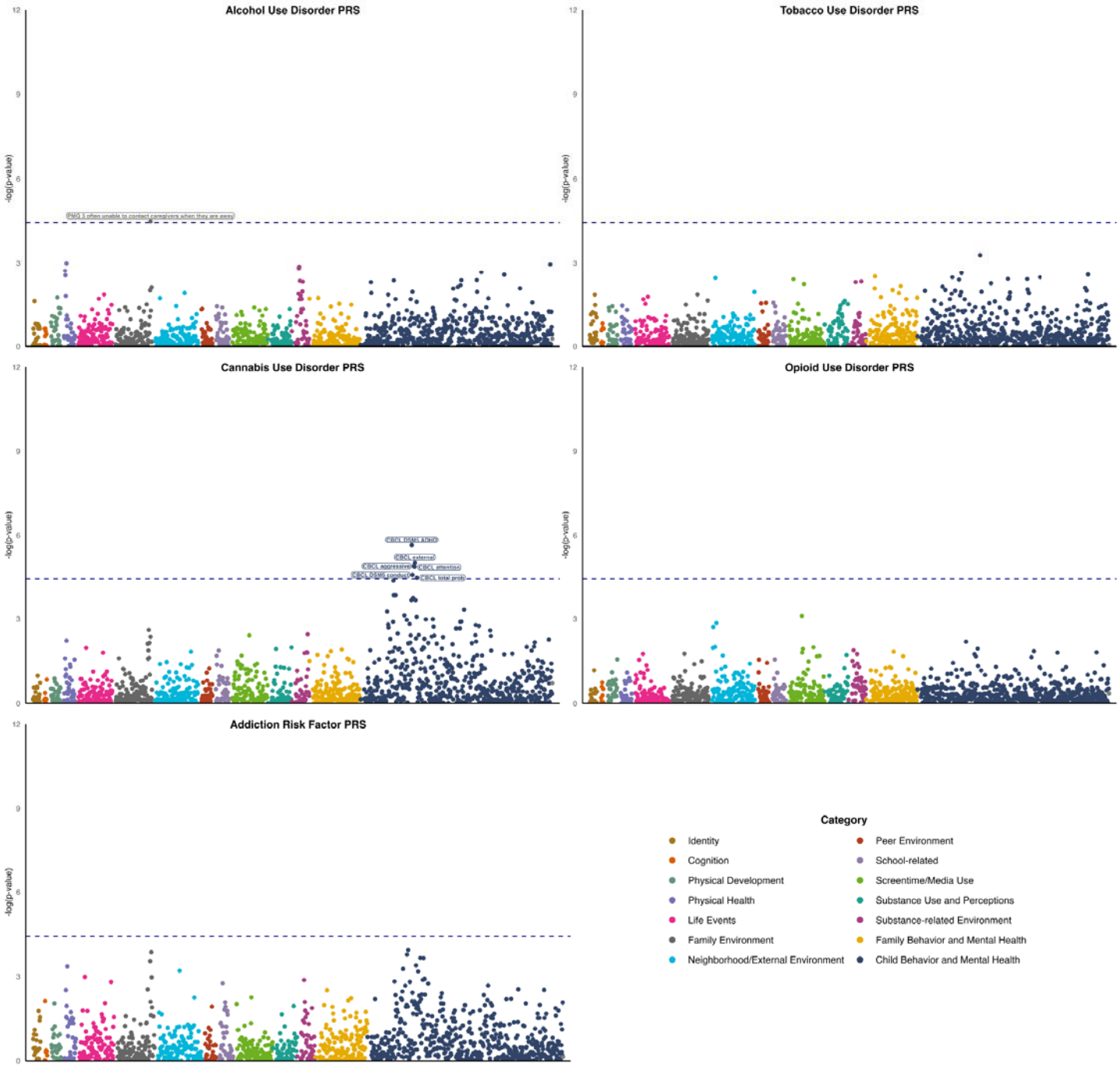
Note. Within-family analyses in European-like genetic ancestry. PRS were decomposed into between- and within-family variance and entered into mixed effects models simultaneously as predictors of phenotypes that were significantly associated with the PRS following Bonferroni correction. Depicted here are the Bonferroni-corrected significant within-family PRS associations (in the darker color), alongside the respective between-family PRS effect (lighter color). Point estimates reflect standardized betas, and error bars represent 95% confidence intervals.

Extended Data Figure 6.

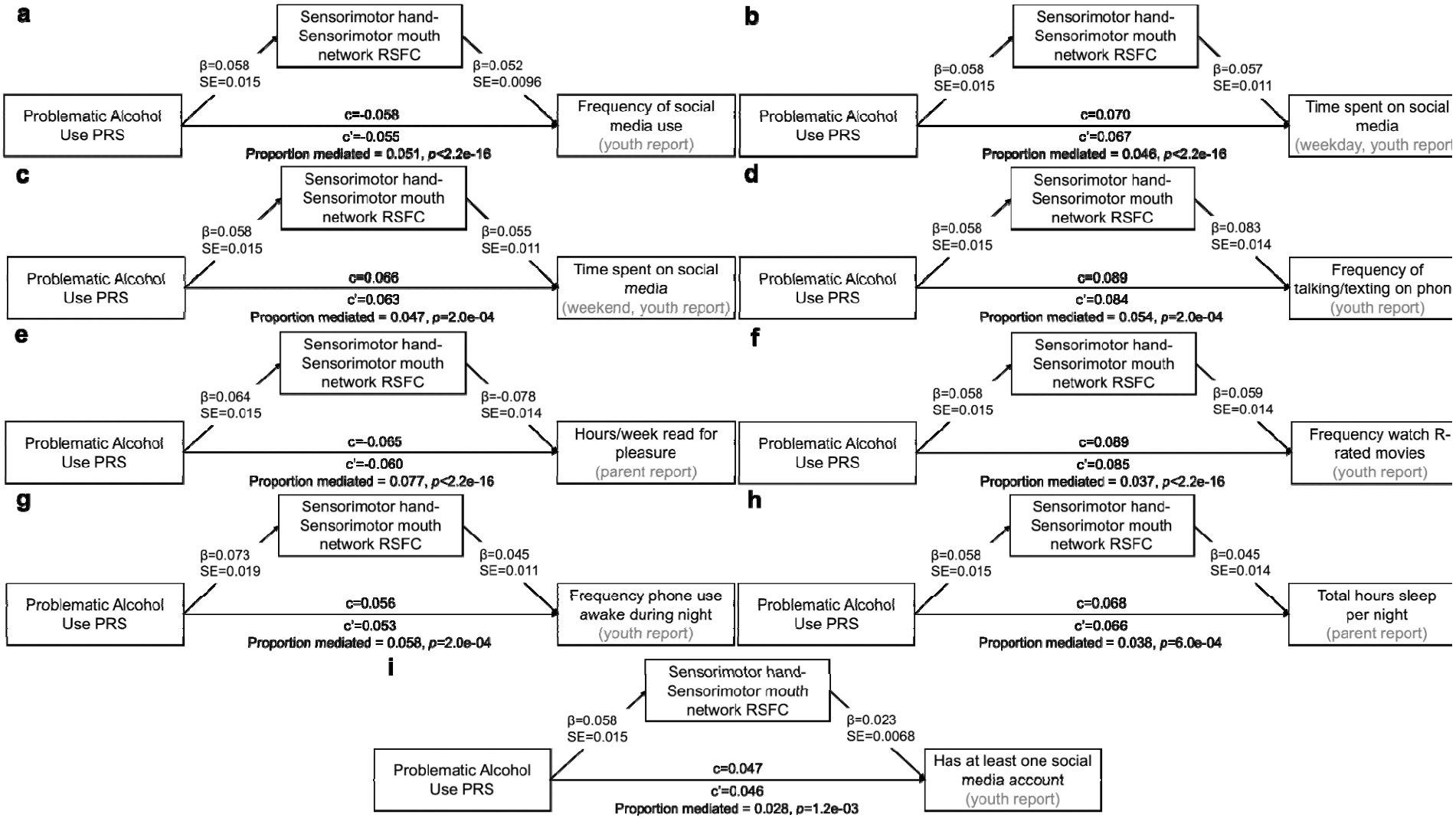


Note. Manhattan plots depicting primary PheWAS associations at the baseline wave among individuals genetically similar to African reference populations. The 940 examined phenotypes are spread along the x-axis, and the y-axes represent the $-\log_{10}$ of the uncorrected *p* values of associations between each PRS and each phenotype. The points are color-coded by phenotype domain. The dashed lines indicate the Bonferroni-significant threshold (i.e., $-\log_{10}(0.05/940) = 4.27$). Full results are located in **Supplemental Table 21**.

Extended Data Figure 7.

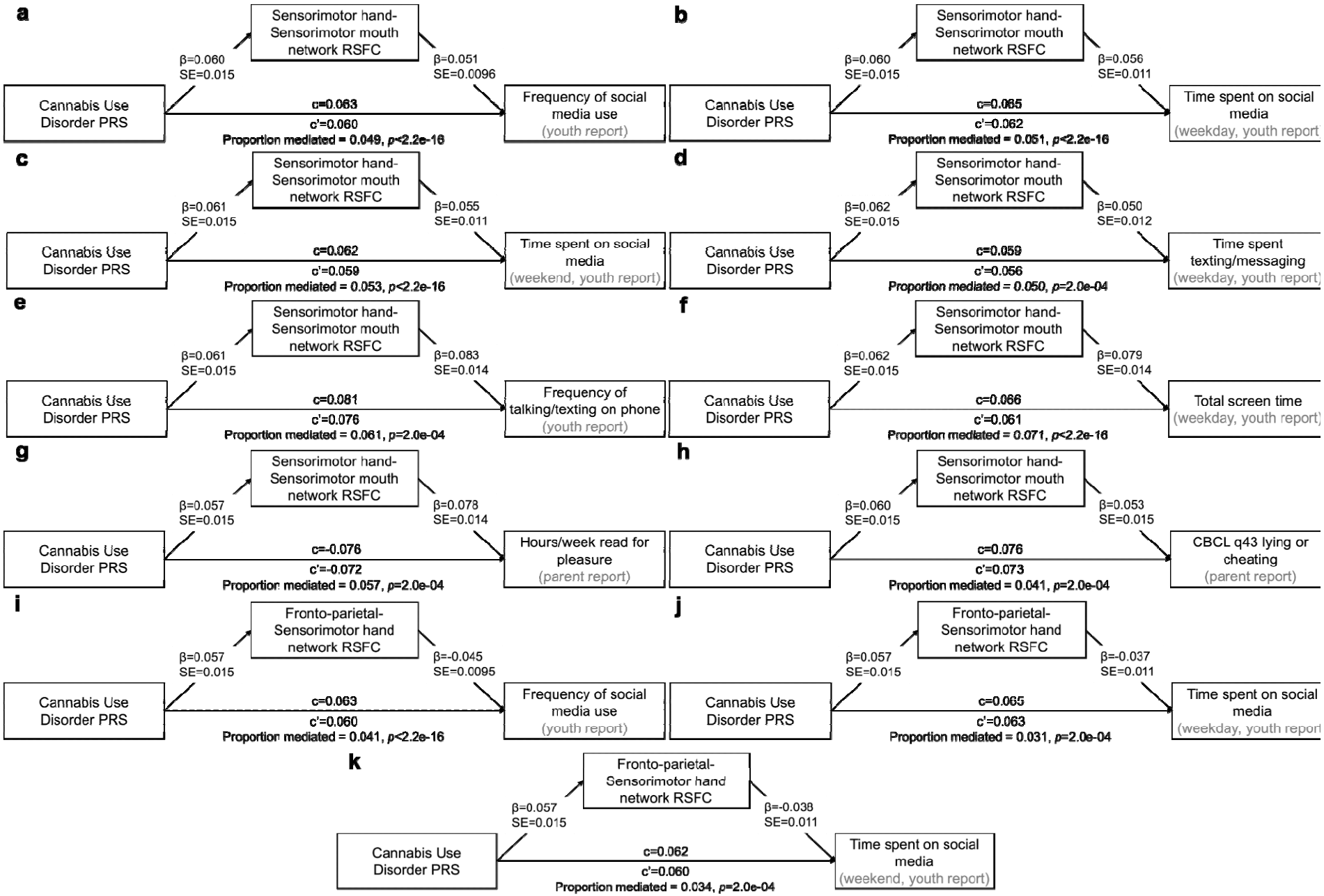


Note. Manhattan plots depicting primary PheWAS associations at the two-year follow-up wave among individuals genetically similar to African reference populations. The 1,357 examined phenotypes are spread along the x-axis, and the y-axes represent the -$\log_{10}$ of the uncorrected *p* values of associations between each PRS and each phenotype. The points are color-coded by phenotype domain. The dashed lines indicate the Bonferroni-significant threshold (i.e., -$\log_{10}$(0.05/1,357) = 4.43). Full results are located in **Supplemental Table 22**.

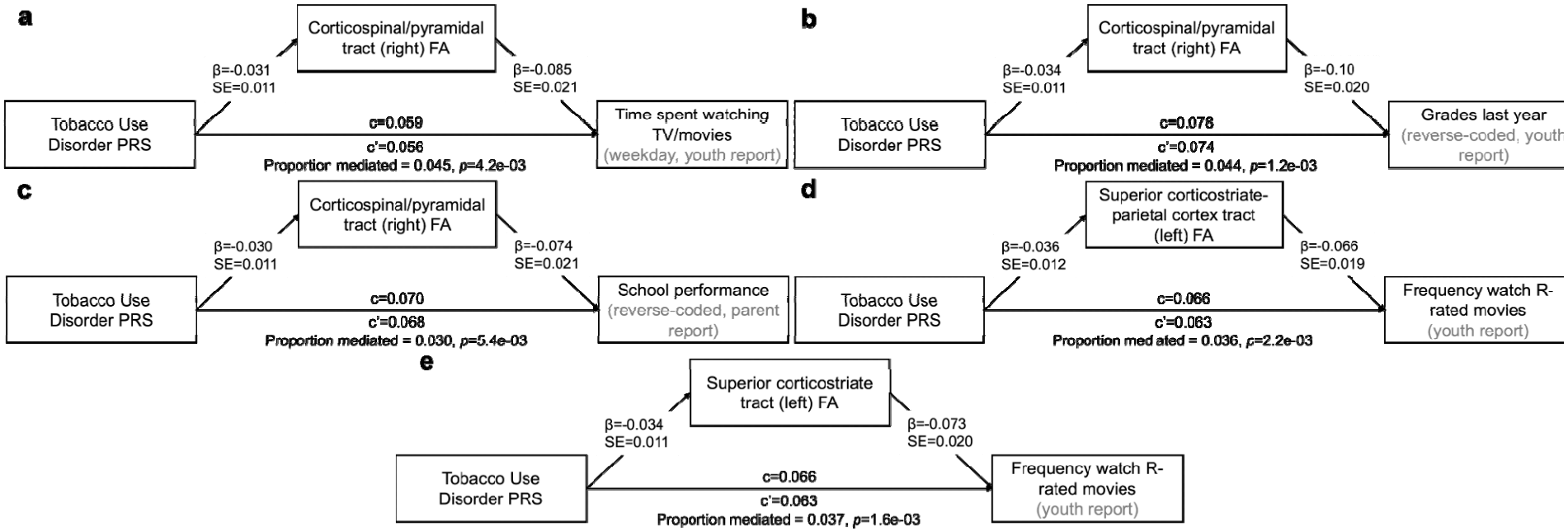


Extended Data Figure 8. Indirect effects of Problematic Alcohol Use PRS on screen time, reading, and sleep phenotypes through resting state functional connectivity (RSFC) between the sensorimotor hand and sensorimotor mouth networks. c = total effect; c' =

direct effect. All indirect effects shown are statistically significant following Bonferroni correction. Results are also shown in **Supplemental Table 35.** These results are specific to individuals genetically similar to European reference populations.


a
Sensorimotor hand-Sensorimotor mouth network RSFC
β=0.060 SE=0.015
β=0.051 SE=0.0096
Cannabis Use Disorder PRS
c=0.063
c'=0.060
Proportion mediated = 0.049, p<2.2e-16
Frequency of social media use (youth report)
b
Sensorimotor hand-Sensorimotor mouth network RSFC
β=0.060 SE=0.015
β=0.056 SE=0.011
Cannabis Use Disorder PRS
c=0.065
c'=0.062
Proportion mediated = 0.051, p<2.2e-16
Time spent on social media (weekday, youth report)
c
Sensorimotor hand-Sensorimotor mouth network RSFC
β=0.061 SE=0.015
β=0.055 SE=0.011
Cannabis Use Disorder PRS
c=0.062
c'=0.059
Proportion mediated = 0.053, p<2.2e-16
Time spent on social media (weekend, youth report)
d
Sensorimotor hand-Sensorimotor mouth network RSFC
β=0.062 SE=0.015
β=0.050 SE=0.012
Cannabis Use Disorder PRS
c=0.059
c'=0.056
Proportion mediated = 0.050, p=2.0e-04
Time spent texting/messaging (weekday, youth report)
e
Sensorimotor hand-Sensorimotor mouth network RSFC
β=0.061 SE=0.015
β=0.083 SE=0.014
Cannabis Use Disorder PRS
c=0.081
c'=0.076
Proportion mediated = 0.061, p=2.0e-04
Frequency of talking/texting on phone (youth report)
f
Sensorimotor hand-Sensorimotor mouth network RSFC
β=0.062 SE=0.015
β=0.079 SE=0.014
Cannabis Use Disorder PRS
c=0.066
c'=0.061
Proportion mediated = 0.071, p<2.2e-16
Total screen time (weekday, youth report)
g
Sensorimotor hand-Sensorimotor mouth network RSFC
β=0.057 SE=0.015
β=0.078 SE=0.014
Cannabis Use Disorder PRS
c=-0.076
c'=-0.072
Proportion mediated = 0.057, p=2.0e-04
Hours/week read for pleasure (parent report)
h
Sensorimotor hand-Sensorimotor mouth network RSFC
β=0.060 SE=0.015
β=0.053 SE=0.015
Cannabis Use Disorder PRS
c=0.076
c'=0.073
Proportion mediated = 0.041, p=2.0e-04
CBCL q43 lying or cheating (parent report)
i
Fronto-parietal-Sensorimotor hand network RSFC
β=0.057 SE=0.015
β=-0.045 SE=0.0095
Cannabis Use Disorder PRS
c=0.063
c'=0.060
Proportion mediated = 0.041, p<2.2e-16
Frequency of social media use (youth report)
j
Fronto-parietal-Sensorimotor hand network RSFC
β=0.057 SE=0.015
β=-0.037 SE=0.011
Cannabis Use Disorder PRS
c=0.065
c'=0.063
Proportion mediated = 0.031, p=2.0e-04
Time spent on social media (weekday, youth report)
k
Fronto-parietal-Sensorimotor hand network RSFC
β=0.057 SE=0.015
β=-0.038 SE=0.011
Cannabis Use Disorder PRS
c=0.062
c'=0.060
Proportion mediated = 0.034, p=2.0e-04
Time spent on social media (weekend, youth report)

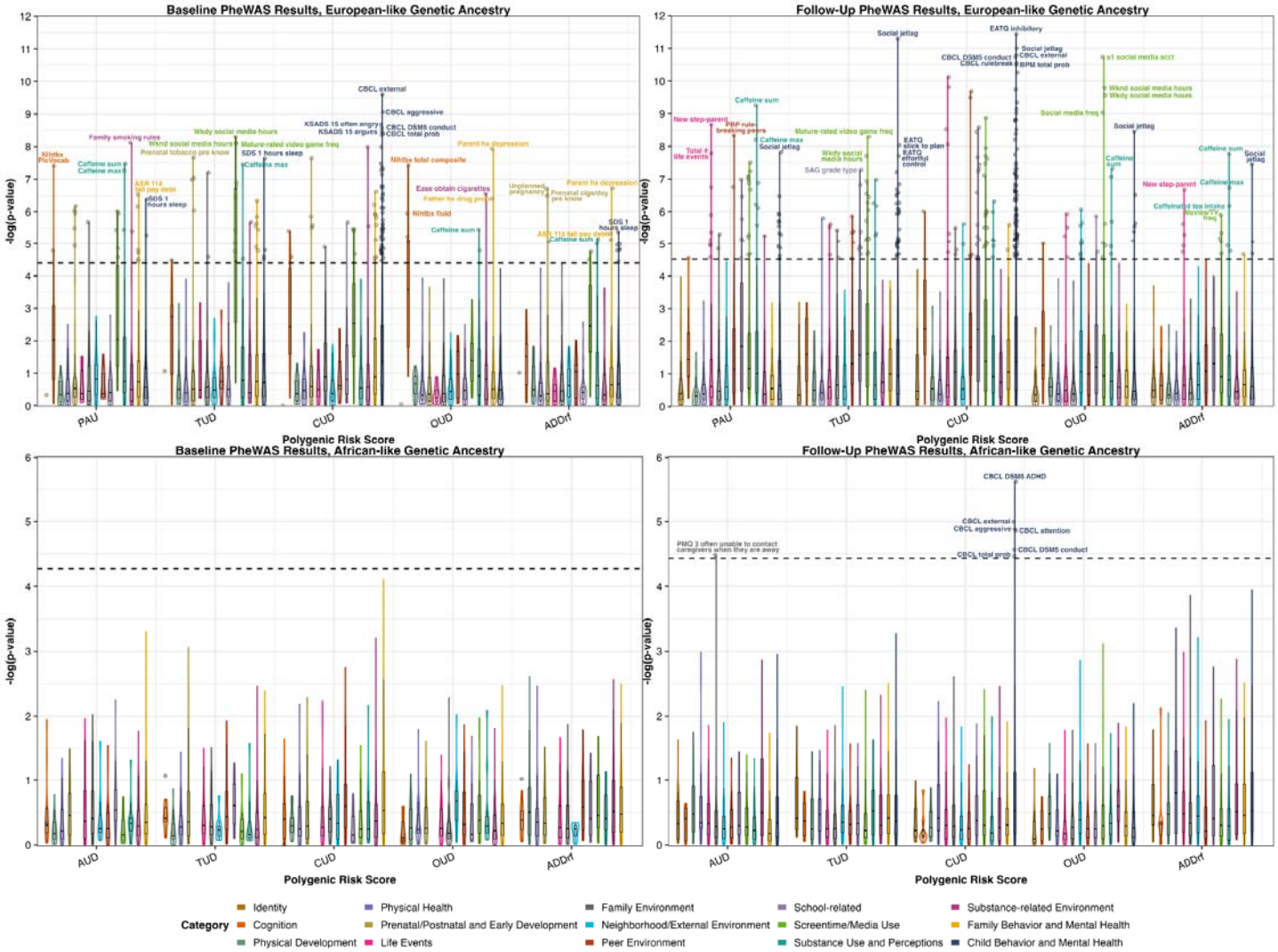
Extended Data Figure 9. Indirect effects of Cannabis Use Disorder PRS on screen time, reading, and lying/cheating phenotypes through resting state functional connectivity (RSFC) between the sensorimotor hand and sensorimotor mouth networks and between the fronto-parietal and sensorimotor hand networks. c = total effect; c' = direct effect. All indirect effects shown are statistically significant following Bonferroni correction. Results are also shown in **Supplemental Table 35.** These results are specific to individuals genetically similar to European reference populations.



Extended Data Figure 10. Indirect effects of Tobacco Use Disorder PRS on screen time and school performance phenotypes through fractional anisotropy in the corticospinal/pyramidal, superior corticostriate-parietal cortex, and superior corticostriate white matter tracts. c = total effect; c' = direct effect. All indirect effects shown are statistically significant following Bonferroni correction. Results are also shown in **Supplemental Table 35.** These results are specific to individuals genetically similar to European reference populations.

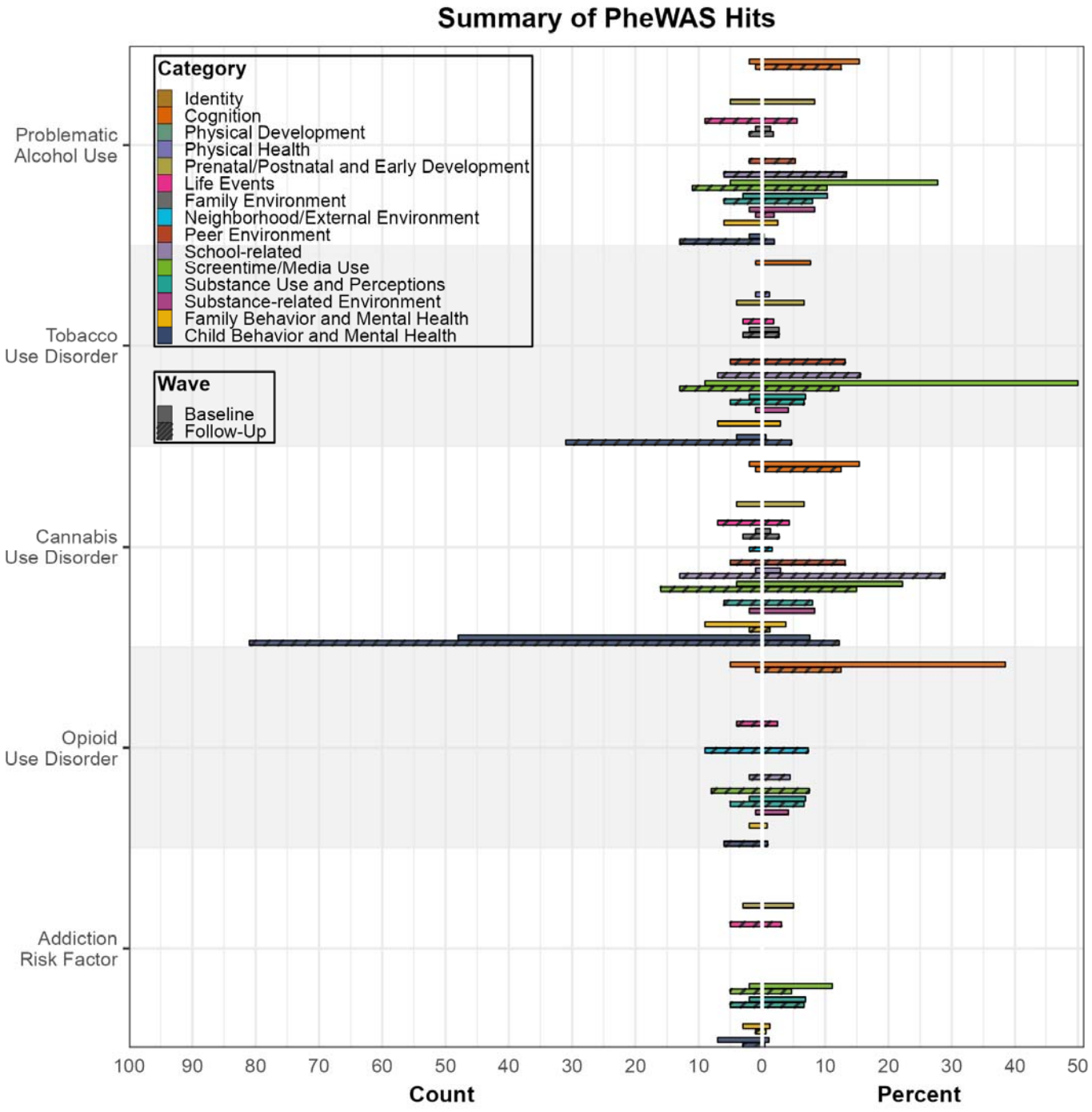
References

Figure 1.



Note. Manhattan and violin plots depicting the number, distribution, and types of phenotypes, within each category, associated with each of the 5 examined PRS. The x-axes are the 5 PRS, the y-axes are the negative log of the p-value of association (higher value indicates stronger association), and color indicates phenotypic Category. The dashed horizontal lines denote the Bonferroni-significant threshold (i.e., 0.05/number of phenotypes). The violin plots represent the distribution of -log(p-values) within each Category, separately for each PRS. Boxplots within each violin plot show the median -log(p-value of association), and the boxes encompass the middle 50% of associations. PAU = Problematic Alcohol Use; TUD = Tobacco Use Disorder; CUD = Cannabis Use Disorder; OUD = Opioid Use Disorder; ADDrf = Addiction Risk Factor; AUD = Alcohol Use Disorder.

Figure 2.



Note. Summary of the number ("Count," left) and percent (right) of associations observed between each PRS (y-axis) and phenotypes in each category, for European-like genetic ancestry. Colors denotes different categories, and solid and dashed bars represent Baseline and Two-Year Follow-Up associations, respectively. Percent reflects the percent of phenotypes in a category associated with a particular PRS of the total number of phenotypes in that category.