ARTICLE   OPEN

Check for updates

Trans-ancestry epigenome-wide association meta-analysis of DNA methylation with lifetime cannabis use

Fang Fang [1✉], Bryan Quach [1], Kaitlyn G. Lawrence[2], Jenny van Dongen[3,4], Jesse A. Marks [1], Sara Lundgren[5], Mingkuan Lin[6], Veronika V. Odintsova [3,4,7], Ricardo Costeira[8], Zongli Xu[2], Linran Zhou[1], Meisha Mandal[1], Yujing Xia [8], Jacqueline M. Vink[9], Laura J. Bierut [10], Miina Ollikainen[5,11], Jack A. Taylor[2], Jordana T. Bell [8], Jaakko Kaprio [5], Dorret I. Boomsma [3,4], Ke Xu [6,12], Dale P. Sandler[2], Dana B. Hancock [1] and Eric O. Johnson [1,13]

Cannabis is widely used worldwide, yet its links to health outcomes are not fully understood. DNA methylation can serve as a mediator to link environmental exposures to health outcomes. We conducted an epigenome-wide association study (EWAS) of peripheral blood-based DNA methylation and lifetime cannabis use (ever vs. never) in a meta-analysis including 9436 participants (7795 European and 1641 African ancestry) from seven cohorts. Accounting for effects of cigarette smoking, our trans-ancestry EWAS meta-analysis revealed four CpG sites significantly associated with lifetime cannabis use at a false discovery rate of 0.05 ($p<5.85\times10^{-7}$): cg22572071 near gene *ADGRF1*, cg15280358 in *ADAM12*, cg00813162 in *ACTN1*, and cg01101459 near *LINC01132*. Additionally, our EWAS analysis in participants who never smoked cigarettes identified another epigenome-wide significant CpG site, cg14237301 annotated to *APOBR*. We used a leave-one-out approach to evaluate methylation scores constructed as a weighted sum of the significant CpGs. The best model can explain 3.79% of the variance in lifetime cannabis use. These findings unravel the DNA methylation changes associated with lifetime cannabis use that are independent of cigarette smoking and may serve as a starting point for further research on the mechanisms through which cannabis exposure impacts health outcomes.

INTRODUCTION

Cannabis use is highly prevalent around the world [1]. In the United States, the legal use of cannabis has expanded across the states over time [2]. Despite the potential therapeutic benefits of medical use [3, 4], the widespread recreational use of cannabis has raised concerns because of its reported associations with many adverse health outcomes, including mental health (anxiety, depression, psychosis, schizophrenia, and mania) [5–8], cognitive deficits [9–11], and addiction [12]. It is a pressing public health issue to better understand the full spectrum of the benefits and adverse consequences associated with cannabis use.

DNA methylation (DNAm), which involves the addition of a methyl group to the C5 position of cytosine in the context of CpG dinucleotides, has been extensively studied in relation to gene expression and can be influenced by the genome, the environment, and stochastic processes [13–15]. The DNAm changes induced by environmental exposure are sometimes persistent and long-lasting, while others are transient and reversible. For example, cigarette smoking has been shown to induce DNAm changes at CpGs throughout the genome. Some of these DNAm changes may revert after smoking cessation, while other DNAm changes may persist for years after cessation [16].

In recent years, research toward understanding the effect of cannabis use on DNAm has grown [17]. Previous candidate gene studies identified DNAm changes of *CB1 receptor* [18] and *DAT1* [19] in cannabis-dependent patients, *COMT* in adolescents defined as high-frequent cannabis users (> four times in the past 4 weeks) [1], and *DRD2* and *NCAM2* in moderate to heavy cannabis users (> 10 days in the last 30 days) [20]. The first epigenome-wide association study (EWAS) of cannabis use, which compared DNAm between 12 cannabis users and 12 non-users in human sperm, found at least 10% DNAm differences at 3979 CpG sites [21]. Our group performed the first blood-based EWAS of lifetime cannabis use (ever vs. never) in 2583 women and found significant DNAm changes at cg15973234 (*CEMIP*) [22]. Cannabis use–associated DNAm changes in blood have also been reported in heavy cannabis users ($N = 96$) [23] and adolescents ($N = 525$) [24]. Taken together, these studies provide evidence that cannabis use impacts the epigenome, but knowledge of specific DNAm changes remain limited.

[1]GenOmics and Translational Research Center, RTI International, Research Triangle Park, NC, USA. [2]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, USA. [3]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. [4]Amsterdam Public Health Research Institute, Amsterdam, The Netherlands. [5]Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland. [6]Department of Psychiatry, Yale School of Medicine, West Haven, CT, USA. [7]Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [8]Department of Twin Research & Genetic Epidemiology, King's College London, London, UK. [9]Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands. [10]Department of Psychiatry, Washington University in Saint Louis School of Medicine, St. Louis, MO, USA. [11]Minerva Foundation Institute for Medical Research, Helsinki, Finland. [12]VA Connecticut Healthcare System, West Haven, CT, USA. [13]Fellow Program, RTI International, Research Triangle Park, NC, USA. ✉email: ffang@rti.org

In this study, we conducted the largest trans-ancestry EWAS meta-analysis for lifetime cannabis use (ever vs. never) in 9436 participants from seven cohorts. The initial model, which adjusted for sex, age at blood collection, blood cell proportions, and technical covariates, yielded 608 significant (False Discovery Rate (FDR) < 0.05) CpGs, among which 82% overlapped with prior EWAS findings for cigarette smoking, whose confounding effects have been found for many other phenotypes like educational attainment [25, 26], aggressive behavior [27] and coffee consumption [28]. We explored this finding further by first adjusting the analyses for cigarette smoking status and next conducting the EWAS in participants who never smoked cigarettes. These two analyses identified a total of five cigarette smoking-independent CpGs significantly associated with lifetime cannabis use. We evaluated these findings by constructing a methylation score, summarizing regional DNAm changes using differential methylation region (DMR) analysis, and integrating the DNAm findings with gene expression and genetic data.

METHODS

Study cohorts

This study included data from seven participating cohorts: the Sister Study [29], Gulf Long-Term Follow-Up Study (GuLF) [30], Netherlands Twin Register (NTR) [31], Veteran Aging Cohort Study (VACS) [32], Finnish Twin Cohort (FinnTwin) [33], Avon Longitudinal Study of Parents and Children (ALSPAC) [34], and UK Adult Twin Registry (TwinsUK) [35]. The final sample size consisted of 9436 participants, including 4190 individuals who reported ever using cannabis and 5246 who reported never using cannabis (Table 1). Detailed information for each cohort can be found in the Supplementary Methods. Informed consent was obtained from each participant, and each study was approved by their Institutional Review Boards.

Cannabis assessment

Our analyses focused on lifetime cannabis use based on self- or parent-report. Participants were classified as ever users if they reported using cannabis at least once prior to the blood sample collection used to generate DNAm data, and as never users if they reported never using cannabis prior to the blood draw. This definition of the phenotype aligns with the "Substances—Lifetime Use" variable in the PhenX Toolkit [36], making the results comparable and combinable.

DNA methylation measurements

DNAm was measured in peripheral blood using either the Illumina Infinium HumanMethylation450 BeadChip (450K array, 76%) or the Illumina Infinium Methylation EPIC BeadChip (EPIC array, 24%), as shown in Table 1. DNAm levels were calculated as β-values, which represent the percentage of DNA that is methylated at the interrogated CpG site and ranges from 0 to 1. Quality control and normalization procedures were implemented consistently across all cohorts, with considerations specific to each cohort (detailed in the Supplementary Methods). Cohort-specific methods are shown in Supplementary Table 1.

EWAS for lifetime cannabis use

In each cohort, the association between DNAm levels and lifetime cannabis use was tested under a linear model or a GEE model if participants were related. We stratified the EWAS analyses by genetic ancestry groups (EA and AA) and DNAm array types (450K and EPIC). For each CpG site, the DNAm beta value was considered as the outcome with lifetime cannabis use as the predictor of interest, and two separate models were applied. In the basic model (Model 1), we included sex (except in cohorts with only one sex), age at blood collection, blood cell type estimation, and technical covariates. In Model 2, we additionally adjusted for cigarette smoking status defined as current, former, or never. Comparing with prior smoking EWAS [37], the results from Model 2 were considered as cigarette smoking-independent DNAm biomarkers for lifetime cannabis use. Additionally, we implemented EWAS within the subset of participants who never smoked cigarettes using Model 1 to minimize the possible confounding effect of cigarette smoking. More detailed information for models and covariates used in each cohort is provided in the Supplementary Methods.

Table 1. Characteristics of the cohort participants.

Study cohort	Total N	Countries	Ancestry	Female %	Age at blood draw[a]	DNAm array	Cannabis ever users%	Never smokers	Cannabis ever use in never smokers%
Sister study	2583	US	100% EA	100%	57.0 ± 8.8	450K	49%	1356	38%
GuLF	1195	US	56% EA 44% AA	0%	44.3 ± 11.8	EPIC	56%	594	40%
NTR	2142	Netherland	Primary EUR[b]	70%	37.4 ± 13.2	450 K	26%	1253	16%
VACS	1116	US	100% AA	0%	48.9 ± 7.9	450K,EPIC	79%	191	54%
Finntwin	1368	Finland	100% EUR	55%	23.5 ± 1.9	450K,EPIC	27%	467	12%
ALSPAC	922	UK	100% EUR	52%	17.1 ± 1.0	450K	40%	218	9%
TwinsUK	110	UK	100% EUR	100%	58.8 ± 9.0	EPIC	27%	67	12%

EA European American, EUR European, AA African American.
[a]Mean and standard deviation presented.
[b]The majority of NTR have Dutch ancestry and individuals who are outliers on genetic PCs were removed.

Meta-analysis

We summarized cohort- and ancestry-specific EWAS results using inverse variance fixed effects meta-analysis implemented in the METAL software [38], with Model 1 (all participants), Model 2 (all participants), and Model 1 in participants who never smoked cigarettes, respectively. We reported overlapping CpGs between 450K and EPIC arrays, which included 452,453 CpG probes. Epigenome-wide significance was defined as FDR less than 5%. Manhattan and QQ plots were geneated using the *CMplot* function within the R package rMVP [39]. Heterogeneity among the studies was assessed using the Cochran's Q-test implemented in METAL.

Methylation scores

As the single largest cohort included in this study, the Sister Study was reserved as the testing dataset to evaluate DNAm scores. For each individual in the Sister Study, a methylation score was calculated as a weighted sum of CpGs significantly associated with lifetime cannabis use in the EWAS meta-analysis conducted without the Sister Study [40]. At a given CpG site i, the methylation beta value ($meth_i$) was multiplied by the effect size of the CpG in the meta-analysis (eff_i). Then a methylation score was obtained by summing over a selected CpG set:

$$methylation\ score = meth_1 \cdot eff_1 + meth_2 \cdot eff_2 + \cdots + meth_n \cdot eff_n$$

We applied different p-value thresholds to select the significant CpG sets for both Model 1 ($p<10^{-1}, 10^{-3}, 10^{-5}, 10^{-7}, 10^{-9}, 10^{-11}, 10^{-13}$) and model 2 ($p<10^{-1}, 10^{-3}, 10^{-5}, 10^{-7}$). To evaluate performance of the methylation scores, we quantified the percentage of variance (R^2) explained as proposed by Lee et al. [41] for binary responses.

Integrating EWAS results with gene expression

To investigate the potential relationship between DNAm and gene expression levels, we used the correlations between DNAm and expression data available in six tissues (brain, colon, kidney, liver, stomach, testis) in the EWAS Atlas [42]. Specifically, we analyzed the DNAm levels of the significant CpGs to determine if they were associated with the expression levels of nearby genes.

Correlation of DNAm between blood and brain tissues

An online database (https://epigenetics.essex.ac.uk/bloodbrain/) [43] was used to examine the correlations of DNAm between whole blood and four brain regions (prefrontal cortex, entorhinal cortex, superior temporal gyrus, and cerebellum), respectively. For each CpG site, a boxplot was generated to display the distribution of DNAm levels across all five tissues, and the Pearson correlation was calculated between the DNAm level in whole blood and each of the four brain tissues.

Enrichment analysis against previous EWAS results

To determine the potential overlap of our top CpGs with previously reported EWAS results, we used the EWAS Atlas toolkit [42] to conduct enrichment analyses. To meet the minimum input requirement, we selected the top 20 CpGs from the EWAS meta-analysis using Model 2 and the top 20 CpGs in participants who never smoked cigarettes, and ran enrichment analyses on each group separately.

Methylation quantitative trait loci (meQTL)

To explore the potential genetic basis for DNAm changes identified in our study, we looked up genetic variants that are associated with DNAm levels of CpG sites within the Genetics of DNA Methylation Consortium (GoDMC) database [44], which includes both local (*cis*) and distal (*trans*) meQTLs. We examined the overlap between the meQTL variants for cannabis use–associated CpG sites and the variants identified in a GWAS for lifetime cannabis use [45].

DMR analyses

We used an R software tool, ipDMR [46], to identify DMRs in which a cluster of correlated CpGs showed evidence for association with lifetime cannabis use. ipDMR calculates an overall p-value for small intervals bordered by two adjacent CpGs based on the association p-values from an EWAS analysis. It then combines all nearby significant intervals (using a seed threshold) and calculates an FDR-adjusted p-value for the combined region. We applied ipDMR to summary statistics from our EWAS meta-analysis with the following parameters: seed p-value < 0.05, maximum

distance to combine adjacent intervals 1000 bp and bin size 50 bp. To assess the biological significance of the identified DMRs, we then conducted gene set enrichment analysis using the tool provided by Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA) [47] to test for enrichment of the genes that overlapped with the DMRs in predefined pathways.

Sensitivity analyses

To evaluate the potential confounding effects of alcohol use and BMI on the association between DNAm and lifetime cannabis use, we conducted sensitivity analyses. In these analyses, we compared the effect sizes and p-values of the significant CpGs identified in the main analysis with those obtained in the sensitivity analyses to assess the robustness of our results. Detailed definitions of alcohol use in each cohort can be found in the Supplementary Methods section.

Comparison of results across ancestry groups

Considering the differential DNAm patterns across different ancestry groups, we conducted a separate EWAS meta-analysis in EA ($N = 7795$) and AA ($N = 1641$) groups with Model 2. We compared the effect sizes and p-values of the top significant CpGs from the main EWAS results with the results in each ancestry group.

RESULTS

Sample characteristics

The demographic characteristics of 9436 study participants are summarized in Table 1. The sample consisted of 57% females, and the mean age at DNAm sampling ranged from 17.1 years in the ALSPAC cohort to 58.8 years in the TwinsUK cohort. A total of 44% of the participants reported having used cannabis at some point in their lives. Within the subset of participants who reported never smoking ($N = 4146$), 27% indicated having used cannabis at least once. DNAm sites were assessed using either the Illumina HumanMethylation450K BeadChip (76%) or the Illumina Human-Methylation EPIC (850K) BeadChip (24%). The sample included individuals of both European ancestry (EA, 83%) and African ancestry (AA, 17%).

EWAS meta-analysis results for lifetime cannabis use

We conducted EWAS meta-analysis on peripheral blood DNAm data from 9436 participants for lifetime cannabis use. In each cohort, we tested the association between lifetime cannabis use and DNAm at each CpG site using either a linear model (for unrelated participants) or a generalized estimation equation (GEE) model (for related participants) with the family ID as the clustering variable. To investigate the confounding effect of cigarette smoking, we compared two models. Model 1 included sex (in cohorts with more than one sex), age at blood collection, measured or estimated white blood cell proportions, and technical covariates. In addition to these covariates, Model 2 included cigarette smoking status defined as current, former, or never. The EWAS meta-analysis with Model 1 identified 608 CpGs significantly associated with lifetime cannabis use at an FDR threshold of 5%. Of these, 500 CpGs had been previously reported as being significantly associated with cigarette smoking [37, 48] (Supplementary Fig. 1). After adjusting for cigarette smoking, Model 2 EWAS meta-analysis identified four CpGs (Fig. 1; Table 2) significantly (FDR < 0.05) associated with lifetime cannabis use. None of these four cannabis use–associated CpGs had been reported as being significant in previous EWAS for cigarette smoking after accounting for multiple testing ($p>0.05/4$) [48]. The quantile-quantile (QQ) plots from both models suggested minimal inflation ($\lambda = 1.1$). Although many CpGs did not reach epigenome-wide significance with Model 2, their effect sizes showed consistent directions of associations as in Model 1 (Supplementary Fig. 2), and their effect sizes were highly correlated ($r = 0.85$). The four cigarette smoking-independent CpGs identified with Model 2—cg01101459, cg22572071, cg15280538, and cg00813162—
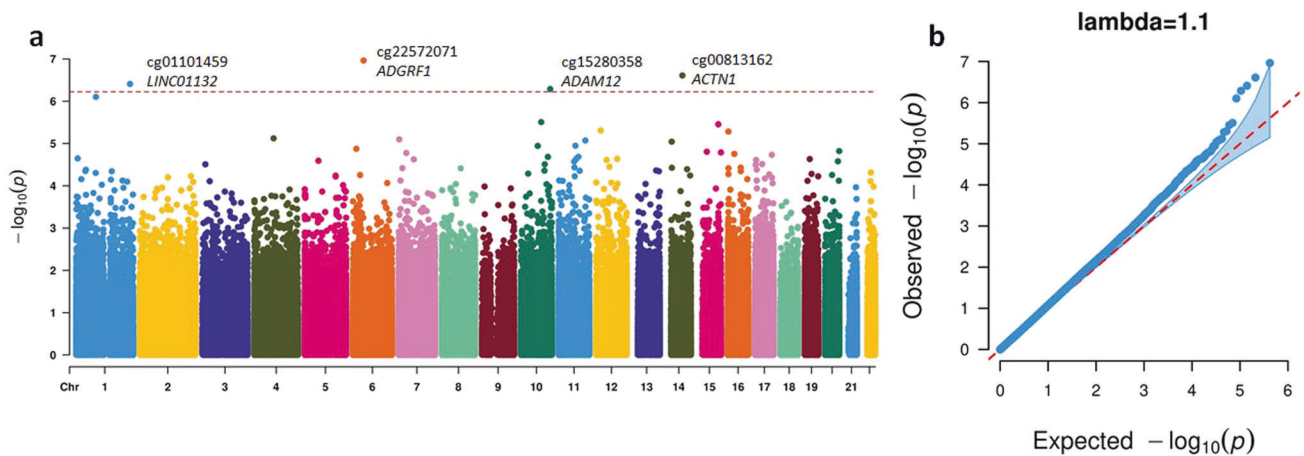
Fig. 1 Results from the EWAS meta-analysis for lifetime cannabis use with Model 2 adjusted for cigarette smoking. a Manhattan plot. The dotted red line indicates the epigenome-wide significance cutoff at FDR < 0.05 ($P < 5.96 \times 10^{-7}$). **b** QQ plot.

were annotated to *LINC01132*, *ADGRF1*, *ADAM12*, and *ACTN1*, respectively, as the closest genes (Supplementary Fig. 3). The full list of top CpGs ($p<0.001$) from the Model 2 EWAS meta-analysis are summarized in Supplementary Table 2, and a forest plot showing the effect sizes of the four significant CpGs is in Supplementary Fig. 4.

Considering potential differences by ancestry, we performed EWAS meta-analyses stratified into EA ($N = 7795$) and AA ($N = 1641$) groups. The ancestry-specific results for the top CpGs reported from the Model 2 EWAS meta-analysis are shown in Supplementary Table 2. The effect sizes for the top CpGs were highly correlated between the EWAS results in the EA and AA groups ($r = 0.77$, Supplementary Fig. 5).

We additionally compared EWAS meta-analyses results within male-only cohorts (VACS and GuLF, $N = 2311$) and female-only cohorts (Sister Study and TwinsUK, $N = 2693$). More than 95% of the top CpGs demonstrated consistent effect directions between males and females (Supplementary Table 2). The correlation of the effect sizes of the top CpGs was 0.69 between males and females (Supplementary Fig. 6).

EWAS meta-analysis results for lifetime cannabis use in participants who never smoked cigarettes
To further explore the confounding effect of cigarette smoking on DNAm, we conducted an EWAS meta-analysis on the subset of participants who reported never having smoked cigarettes ($N = 4146$). The EWAS meta-analysis, using Model 1, identified one CpG significantly associated with lifetime cannabis use at FDR < 0.05 in this subset of participants (Fig. 2; Table 2). This CpG site, cg14237301, is annotated to the gene *APOBR*, which has been reported to be significantly associated with lifetime cannabis use in a genome-wide association study (GWAS) [45] (Fig. 3). The full list of top CpGs ($p<0.001$) from the EWAS meta-analysis in participants who never smoked cigarettes is summarized in Supplementary Table 3. Their effect sizes were highly correlated with the EWAS meta-analysis results in all participants under either model ($r>0.7$, Supplementary Fig. 2).

Replication of previous reported DNAm changes associated with cannabis use
We compared our results with previously reported EWAS for cannabis use (Supplementary Table 4). All seven nominally significant CpGs identified in Markunas et al. [49] remained significant in our EWAS results using Model 2 after applying a Bonferroni correction ($0.05/7 = 0.0071$). Osborne et al. reported CpGs associated with cannabis use in individuals who exclusively used cannabis or used it in conjunction with tobacco [23]. We

compared them with our findings in never smokers and EWAS model 1, separately. Out of a total of 30 CpGs examined, we successfully replicated 10 CpGs in our study with a Bonferroni correction. Additionally, a recent study examined the associations between DNAm and recent and cumulative cannabis use in middle-aged adults [50]. Our results replicated four out of 40 CpGs reported with a Bonferroni correction.

Methylation scores
To assess the ability of DNAm levels to predict lifetime cannabis use, we calculated a methylation score based on summary statistics from the EWAS meta-analyses without the Sister Study, the largest cohort in this project. We then compared the performance of multiple models with different p-value cutoffs by the variance in lifetime cannabis use explained by the methylation scores (Supplementary Table 5). The best-performing methylation score, based on 50 CpGs with $p<10^{-9}$ in the EWAS meta-analysis using Model 1 (Supplementary Table 6), explained 3.79% of the variance of lifetime cannabis use in the Sister Study ($p = 1.71 \times 10^{-17}$). To further examine the confounding effect of cigarette smoking, we also evaluated the variance explained by the methylation score in participants who never smoked cigarettes. In this subset of participants, the methylation score based on the same 50 CpGs explained 0.91% of the variance of lifetime cannabis use ($p = 1.19 \times 10^{-3}$). In contrast, the methylation score constructed based on EWAS meta-analysis results with Model 2 can explain 0.58% of the variance of lifetime cannabis use in all participants of the Sister Study, achieved with p-value cutoff at 10^{-5}.

Sensitivity analysis
We conducted additional analyses to adjust for other factors that influence DNAm as shown in previous studies: alcohol use [51] and BMI [52]. We examined if the EWAS results with Model 2 differed by expanding Model 2 to also include alcohol use and body mass index (BMI) as covariates. The effect sizes of the top cannabis use–associated CpGs ($p<0.001$) from Model 2 (Supplementary Table 2) were strongly correlated ($r = 0.99$) with the results of the sensitivity analyses (Supplementary Fig. 7).

Follow-up of CpGs significantly associated with lifetime cannabis use
We examined the five CpGs that were epigenome-wide significantly associated with lifetime cannabis use, four of which were identified using Model 2 in all participants (adjusted for cigarette smoking) and one CpG identified in participants who never smoked cigarettes. Using the EWAS Atlas [42] (https://

Table 2. Epigenome-wide significant (FDR < 0.05) CpGs identified in EWAS meta-analyses for lifetime cannabis use from Model 2, adjusted for cigarette smoking.

CpG	Chr:Position (hg19)	Closest genes	Location[a]	Model 2 in all participants (N = 9436)			Model 1 in all participants (N = 9436)		Model 1 in participants who never smoked cigarettes (N = 4146)		EWAS meta-analysis of cigarette smoking [48] (current vs. non-current)		
				Direction	Effect[b]	P-value	Effect[b]	P-value	Effect[b]	P-value	Effect[c]	Direction	P-value
cg01101459	1:234871477	LINC01132	TFBS	-+---+	−0.0042	3.90E-07	−0.0030	2.27E-04	−0.0034	0.0064	0.18	+++-+	0.014
cg22572071	6:47074382	ADGRF1	Enhancer	+++++ ++++	0.0040	1.09E-07	0.0028	9.51E-05	0.0024	0.025	0.0074	+-+--+	0.91
cg15280358	10:127904472	ADAM12	Gene body	+-++?+ +++-	0.0058	5.15E-07	0.0044	8.32E-05	0.0040	0.016	−0.11	+----+	0.22
cg00813162	14:69443362	ACTN1	5'UTR	-+--- ----	−0.0028	2.46E-07	−0.0024	5.51E-06	−0.0025	0.0019	0.023	++-?+	0.69
cg14237301	16:28506477	APOBR	1st Exon; 5' UTR	----- -?-	−0.0040	1.76e-05	−0.0032	3.10E-4	−0.0071	9.05E-08	0.070	-++++	0.37

[a]TFBS: Transcription Factor Binding Site; 5'UTR: 5' untranslated region.
[b]Positive effect size means higher methylation (hypermethylation) in ever cannabis users.
[c]Positive effect size means higher methylation (hypermethylation) in current cigarette smokers.

ngdc.cncb.ac.cn/ewas/tools), we found that three of the five CpGs were significantly correlated with nearby gene expression in brain (Supplementary Table 7, $cor(cg01101459, LINC01132) = 0.121$, $cor(cg00813162, ACTN1) = −0.135$, $cor(cg14237301, APOBR) = −0.389$).

We also looked up the correlation of DNAm for the five CpGs between whole blood and four brain regions (prefrontal cortex, entorhinal cortex, superior temporal gyrus, and cerebellum) through an existing database (http://epigenetics.essex.ac.uk/bloodbrain/) [43]. No significant correlations were observed after applying the Bonferroni correction. Among the five CpGs examined, moderate correlations ($r > 0.2$) were observed between whole blood and the prefrontal cortex for DNAm at cg22572071-ADGRF1 and cg14237301-APOBR. Similarly, at cg22572071-ADGRF1 and cg15280358-ADAM12, moderate correlations ($r > 0.2$) were observed between whole blood and the cerebellum (Supplementary Fig. 8, Supplementary Table 8).

Enrichment analyses comparing the top cannabis use–associated CpGs from Model 2 to previously reported EWAS results in EWAS Atlas identified associations with other diseases and environment exposures (Supplementary Table 9). Among the enriched traits detected, Crohn's disease, alcohol consumption, BMI, and multiple sclerosis were significantly overlapped with our EWAS results. Interestingly, cannabis use–associated CpGs identified in participants who never smoked cigarettes showed significant overlap with CpGs previously associated with smoking ($p = 5.58 \times 10^{-8}$), smoking cessation, lung function, and lung carcinoma.

To explore whether the DNAm changes were genetic-driven, we looked up the methylation quantitative trait loci (meQTLs) for the five epigenome-wide significant CpGs in the Genetics of DNA Methylation Consortium (GoDMC) database [44] (http://mqtldb.godmc.org.uk/about), and overlapped them with the GWAS results for lifetime cannabis use [45] (Supplementary Table 10). None of the meQTLs were significantly associated with lifetime cannabis use ($p = 5 \times 10^{-8}$), indicating that our significantly associated CpGs were not directly driven by the genetic variants. The meQTLs for cg00813162-ACTN1 and cg14237301-APOBR showed some degree of significance in the GWAS ($p<0.01$, Supplementary Table 10).

Differentially methylated regions

We used an R software tool, ipDMR [46], to identify DMRs based on EWAS meta-analysis results (Supplementary Table 11). Model 1 (no adjustment for cigarette smoking) yielded 514 significant DMRs (FDR < 0.05) with a minimum of two probes, while Model 2 (accounting for cigarette smoking) had 10 significant DMRs, showing that one of the four epigenome-wide significant CpGs (cg00813162-ACTN1) reside in a region where proximate CpGs were associated with lifetime cannabis use. Additionally, there were nine DMRs that included no single epigenome-wide significant CpG, and instead multiple correlated CpGs showed some evidence of association. The DMR analysis in participants who never smoked cigarettes identified six DMRs with a minimum of two probes, none of which included the epigenome-wide significant CpG sites. The gene set enrichment analysis of the genes that overlap with the DMRs from Model 2 identified four Gene Ontology (GO) biological processes (Supplementary Fig. 9): growth, response to metal ion, actin filament bundle organization, and cellular response to zinc ion.

DISCUSSION

In this study, we performed the largest EWAS meta-analyses of lifetime cannabis use to date (9436 multi-ancestry participants) using DNAm data from peripheral blood samples. The basic model showed that DNAm changes associated with lifetime cannabis use largely overlapped with cigarette smoking–associated DNAm sites.
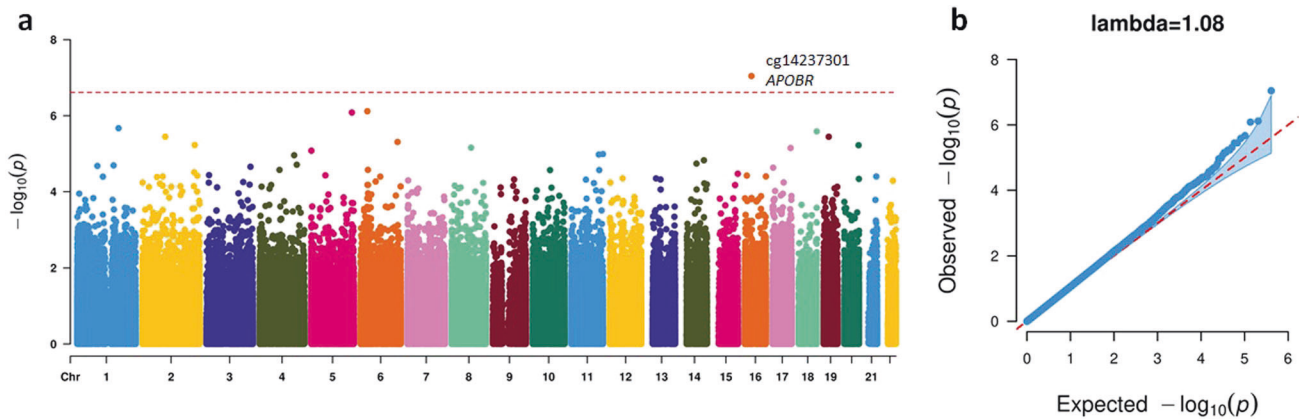
Fig. 2 **Results from the EWAS meta-analysis for lifetime cannabis use in participants who never smoked cigarettes. a** Manhattan plot. The dotted red line indicates the epigenome-wide significance cutoff at FDR < 0.05 ($P < 2.45 \times 10^{-7}$). **b** QQ plot.

After additionally adjusting for smoking in the EWAS model, we found four CpG sites statistically independent of cigarette smoking that were significantly associated with lifetime cannabis use. Additionally, we conducted EWAS in participants who never smoked cigarettes to further eliminate the influence of the participants' cigarette smoking. This analysis showed high consistency with the smoking-adjusted model in all participants (Model 2) and yielded one additional CpG significantly associated with lifetime cannabis use. The genes annotated to these five cannabis use–associated CpGs are relevant to a range of health outcomes [53–63].

The five epigenome-wide significant CpGs are annotated to the nearest genes: *LINC01132*, *ADGRF1*, *ADAM12*, *ACTN1*, and *APOBR*. Of these, cg01101459-*LINC01132*, cg00813162-*ACTN1*, and cg14237301-*APOBR* were inversely associated with cannabis use, meaning that individuals who had used cannabis had lower DNAm at these CpG sites compared to those who had never used cannabis. *LINC01132* is a long noncoding RNA gene that has been reported to function as an oncogene that relates to the malignant behaviors of cancer cells in hepatocellular carcinoma (HCC) [53, 64, 65] and ovarian cancer [64, 65]. Interestingly, cannabis use has been linked to a reduced incidence rate of HCC [66] and the potential for treating ovarian cancer [67, 68]. *ACTN1* (Alpha-Actinin-1) encodes a non-muscle cytoskeletal protein that binds actin to the cell membrane [54]. Genetic variants and differential expression of the *ACTN1* have been reported in various diseases, including congenital macrothrombocytopenia, Angelman syndrome, Bowen disease, postmenopausal osteoporosis, lupus erythematosus, and COVID-19 [54–57, 69, 70]. There is evidence that cannabinoids can attenuate seizures and EEG abnormalities in Angelman syndrome [71, 72], effective in treating individuals with osteoporosis [73], reduce cell proliferation and induce apoptosis in melanoma cells [74], as well as downregulate the immune response in lupus erythematosus [75]. Notably, *ACTN1* was also reported as a differentially methylated gene in sperms associated with cannabis use [21]. *APOBR* (Apolipoprotein B receptor) encodes a receptor protein that binds to dietary triglyceride-rich lipoproteins. Its genetic variants have been associated with obesity [58, 76], bladder cancer [61], pneumonia [60], allergy [59], and lifetime cannabis use [45]. Relatedly, cannabis use has been associated with a reduced obesity rate [77] but an increased risk for pneumonia [78]. Overall, DNAm may serve as a mediator between cannabis use and its impact on health outcomes.

On the other side, cg22572071-*ADGRF1* and cg15280358-*ADAM12* were positively associated with cannabis use. *ADGRF1* is a receptor gene that that is critical in neurodevelopment and neuroinflammation [62, 79]. The overexpression of *ADGRF1* has

been reported in breast cancer [80]. *ADAM12* encodes protein that involving in cell-cell interaction, muscle development, and neurogenesis. The expression of *ADAM12* has been reported to be upregulated in various tumor cells and is an emerging prognostic biomarker for cancer [81–85]. Genetic variants in *ADAM12* have been associated with neurological diseases such as multiple sclerosis and Alzheimer's disease. Taking together, cannabis use is associated with regulation of genes that function in neurogenesis, neurodevelopment, brain structures and onco-genesis. Further experimental studies are needed to determine the potential effects of changes in DNAm levels on the corresponding gene expressions.

The majority of significant CpGs in EWAS for cannabis use with the basic model (Model 1) overlapped with those identified in the EWAS for cigarette smoking [48]. The number of significant CpGs decreased dramatically after adjusting for smoking status in our Model 2. These findings suggest that cigarette smoking is a strong confounder for cannabis use. The overlap with EWAS results for cigarette smoking were also found for many other phenotypes like educational attainment [25, 26], aggressive behavior [27] and coffee consumption [28]. However, we found that DNAm scores calculated as weighted sums of the beta values of significant CpGs from Model 1 explained a greater amount of variance in lifetime cannabis use than the DNAm scores based on Model 2 (3.79% vs. 0.58%). Even in participants who never smoked cigarettes, the DNAm scores based on Model 1 could explain 0.91% of the variance in lifetime cannabis use. Additionally, the EWAS results in participants who never smoked cigarettes showed enrichment of CpGs associated with cigarette smoking. These results suggest that cannabis use and cigarette smoking may independently influence the DNAm levels of these CpGs, implying the potential influence of common combustible chemicals that are shared between cannabis use and cigarette smoking.

In previous studies, DNAm scores have been used to predict various outcomes with the variance explained ranging from 0.6% in low-density lipoprotein, 2.5% in educational attainment, 12.5% in alcohol use, to 60.9% in cigarette smoking [86]. Such predictors may provide more accurate measurements than self-reported phenotypes and could have clinical applications by relating to health outcomes. We assessed whether blood-based DNAm could predict cannabis use and found that the variance in lifetime cannabis use explained by DNAm scores was (3.79%) moderate and not as accurate as DNAm-based score for cigarette smoking. Future applications that integrate both polygenic risk scores and DNAm scores may improve prediction power [40, 86]. The accumulation of even larger scale GWAS and EWAS for lifetime cannabis use will be necessary to establish reliable biomarkers for clinical purposes.
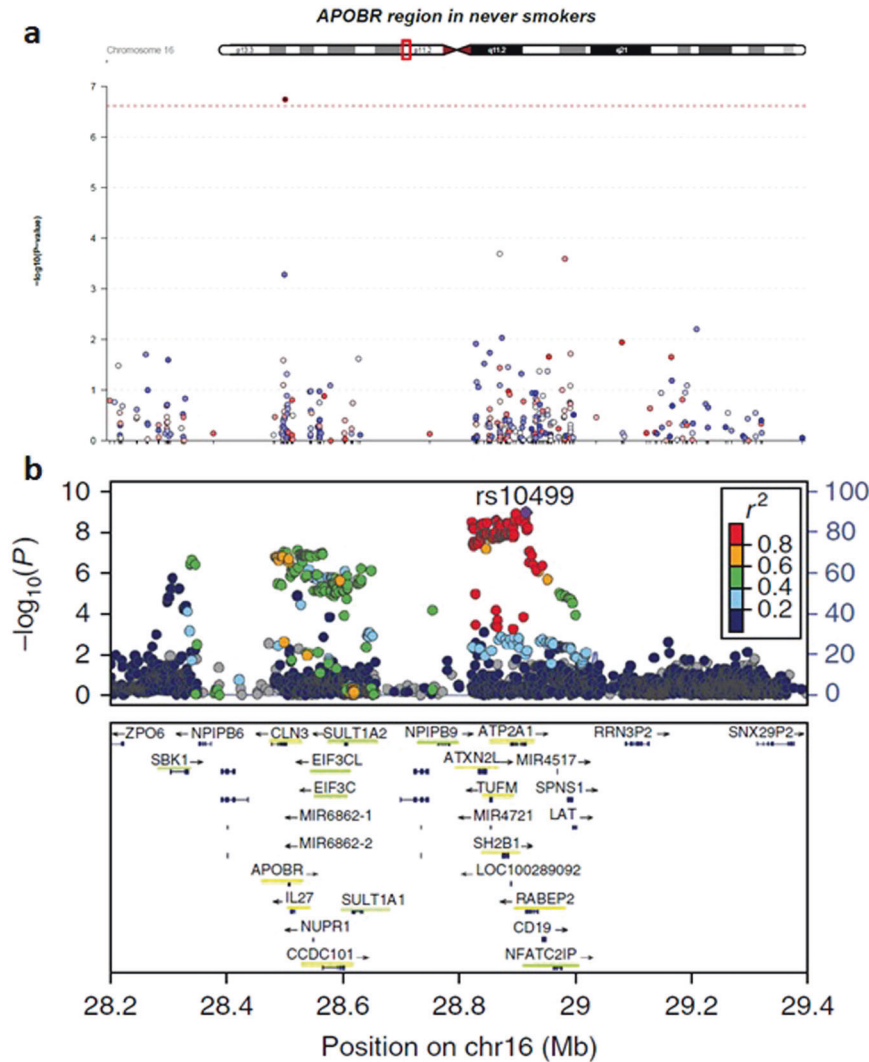
Fig. 3 Regional plot for EWAS results in participants who never smoked cigarettes (a) and GWAS results (b) for lifetime cannabis use around the genes *APOBR/CLN3*. The x-axis shows the genomic position in base pair (bp) in hg19, while the y-axis shows the significance of associations ($-\log_{10}$ p-values). **a** Each dot is a CpG probe, and the red dotted line indicates the epigenome-wide significance at FDR < 0.05. **b** Each dot is a SNP site, and the colors show different levels of LD in r^2. The box in the bottom includes the genes within the genomic region, and genes underlined in yellow were significant in the gene-based test while those underlined in green were identified in the S-PrediXcan analysis [45].

As a complement to the main EWAS, the DMR analysis that combined nearby correlated CpGs identified additional regions associated with cannabis use. After adjusting for smoking, the EWAS for lifetime cannabis use identified fewer epigenome-wide significant hits, and the DMR analysis provides an improved power to detect correlated CpGs with small effects influenced by cannabis exposure. The gene set enrichment analysis revealed biological processes related to growth, response to stress, and assembly of actin filament bundle. The stress response pathways have been linked to the use of Δ^9-tetrahydrocannabinol (THC), which has led to DNA damage and induced oxidative stress in both blood and brain cells [87, 88]. The actin filament bundle gene set was also reported in cellular remodeling events induced by cannabinoid that affect the brain architecture and wiring [89].

The large sample set included in this study provides a good representation for a wide range of populations across countries, sexes, ancestry groups, and age ranges, empowering more generalizable findings in identifying a common and robust DNAm signature for lifetime cannabis use. A recent genetic study [90] has shown that increased sample sizes of diverse ancestries improved detection power and generated more generalizable polygenic risk scores. However, one limitation of this approach is that associations with ancestry effects may have been attenuated when all data were combined, as heterogeneity across different cohorts would have reduced power to detect such specific associations in underrepresented populations. Such heterogeneity is also reflected by inconsistent directions of effects in meta-analysis. Future studies on more data from diverse populations may reveal ancestry-, sex- and age-specific DNAm associations. There may also be relevant confounders that we have not been able to adjust for, such as exposures and experiences that lead to cannabis use.

In this study, we analyzed DNAm profiles from blood samples. While DNAm profiles differ across tissues and cell types, our results based on blood samples may not be generalizable to other tissues that may be more biologically relevant to the addiction and other behavioral effects of cannabis use, such as the brain. It should also be noted that the current EWAS results may suffer from potential confounding effects caused by blood cell subtype heterogeneity that were not fully captured by reference-based deconvolution that we applied [91, 92].

The EWAS meta-analysis conducted in this study only reflects the association between lifetime cannabis use and DNAm levels, without any causal inference. The differentially methylated CpGs identified indicate persistent rather than transient associations with cannabis use. We expect that including ever users who had used cannabis long before blood drawn may bias results towards the null, attenuating effect sizes that may be driven by heavy users within the case group. To delve deeper into the persistent and transient effects of cannabis on DNAm alterations, more precise measurements of cannabis use regarding recency and frequency are needed. Taken together, we propose future studies to integrate cannabis use patterns, GWAS results, DNAm data from multiple tissues, and gene expression data, to infer causal links between cannabis exposure and DNAm levels.

In conclusion, our EWAS found that a large proportion of DNAm changes that are significantly associated with cannabis use overlap with those observed for cigarette smoking, even in participants who never smoked cigarettes, suggesting that cannabis use and cigarette smoking may independently influence the DNAm levels of shared CpGs. After adjusting for smoking status in the EWAS and conducting additional association testing in participants who never smoked cigarettes, we identified five cigarette smoking-independent CpGs that were significantly associated with lifetime cannabis use. The genes associated with these CpGs have been linked to various health outcomes, encompassing both disease risk and potential benefits related to cannabis use, highlighting the role of DNAm in the investigation of cannabis effects. These findings provide insights into DNAm profiles that are shared between smoking and cannabis use or specific to each substance, and suggest a substantial proportion of the variance in lifetime cannabis use are captured by DNAm. Follow-up studies are warranted to unravel the biological relevance of the differential DNAm to health outcomes.

DATA AVAILABILITY
The summary statistics from EWAS meta-analysis and analysis codes are available from the corresponding author upon reasonable request.

REFERENCES
1. Knaap LJ, Schaefer JM, Franken IH, Verhulst FC, Oort FV, Riese H. Catechol-O-methyltransferase gene methylation and substance use in adolescents: the TRAILS study. Genes Brain Behav. 2014;13:618–25.
2. Hasin D, Walsh C. Trends over time in adult cannabis use: a review of recent findings. Curr Opin Psychol. 2021;38:80–5.
3. Wang Y, Hong Y, Yan J, Brown B, Lin X, Zhang X, et al. Low-dose delta-9-tetrahydrocannabinol as beneficial treatment for aged APP/PS1 mice. Int J Mol Sci. 2022;23:2757.
4. Bogale K, Raup-Konsavage W, Dalessio S, Vrana K, Coates MD. Cannabis and cannabis derivatives for abdominal pain management in inflammatory bowel disease. Med Cannabis Cannabinoids. 2021;4:97–106.
5. Radhakrishnan R, Wilkinson ST, D'Souza DC. Gone to pot - a review of the association between cannabis and psychosis. Front Psychiatry. 2014;5:54.
6. Gibbs M, Winsper C, Marwaha S, Gilbert E, Broome M, Singh SP. Cannabis use and mania symptoms: a systematic review and meta-analysis. J Affect Disord. 2015;171:39–47.
7. Malone DT, Hill MN, Rubino T. Adolescent cannabis use and psychosis: epidemiology and neurodevelopmental models. Br J Pharm. 2010;160:511–22.
8. Leweke FM, Koethe D. Cannabis and psychiatric disorders: it is not only addiction. Addict Biol. 2008;13:264–75.
9. Crean RD, Crane NA, Mason BJ. An evidence based review of acute and long-term effects of cannabis use on executive cognitive functions. J Addict Med. 2011;5:1–8.
10. Gilman JM, Schmitt WA, Potter K, Kendzior B, Pachas GN, Hickey S, et al. Identification of 9-tetrahydrocannabinol (THC) impairment using functional brain imaging. Neuropsychopharmacology. 2022;47:944–52.
11. Alegria AA, Hasin DS, Nunes EV, Liu SM, Davies C, Grant BF, et al. Comorbidity of generalized anxiety disorder and substance use disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. J Clin Psychiatry. 2010;71:1187–95.

12. Lopez-Quintero C, Perez de los Cobos J, Hasin DS, Okuda M, Wang S, Grant BF, et al. Probability and predictors of transition from first use to dependence on nicotine, alcohol, cannabis, and cocaine: results of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). Drug Alcohol Depend. 2011;115:120–30.
13. Lowdon RF, Jang HS, Wang T. Evolution of epigenetic regulation in vertebrate genomes. Trends Genet. 2016;32:269–83.
14. Straussman R, Nejman D, Roberts D, Steinfeld I, Blum B, Benvenisty N, et al. Developmental programming of CpG island methylation profiles in the human genome. Nat Struct Mol Biol. 2009;16:564–71.
15. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38:23–38.
16. Kaur G, Begum R, Thota S, Batra S. A systematic review of smoking-related epigenetic alterations. Arch Toxicol. 2019;93:2715–40.
17. Szutorisz H, Hurd YL. Epigenetic effects of cannabis exposure. Biol Psychiatry. 2016;79:586–94.
18. Rotter A, Bayerlein K, Hansbauer M, Weiland J, Sperling W, Kornhuber J, et al. CB1 and CB2 receptor expression and promoter methylation in patients with cannabis dependence. Eur Addict Res. 2013;19:13–20.
19. Grzywacz A, Barczak W, Chmielowiec J, Chmielowiec K, Suchanecka A, Trybek G et al. Contribution of dopamine transporter gene methylation status to cannabis dependency. Brain Sci. 2020;10.
20. Gerra MC, Jayanthi S, Manfredini M, Walther D, Schroeder J, Phillips KA, et al. Gene variants and educational attainment in cannabis use: mediating role of DNA methylation. Transl Psychiatry. 2018;8:23.
21. Murphy SK, Itchon-Ramos N, Visco Z, Huang Z, Grenier C, Schrott R, et al. Cannabinoid exposure and altered DNA methylation in rat and human sperm. Epigenetics. 2018;13:1208–21.
22. Markunas CA, Hancock DB, Xu Z, Quach BC, Fang F, Sandler DP, et al. Epigenome-wide analysis uncovers a blood-based DNA methylation biomarker of lifetime cannabis use. Am J Med Genet B Neuropsychiatr Genet. 2021;186:173–82.
23. Osborne AJ, Pearson JF, Noble AJ, Gemmell NJ, Horwood LJ, Boden JM, et al. Genome-wide DNA methylation analysis of heavy cannabis exposure in a New Zealand longitudinal cohort. Transl Psychiatry. 2020;10:114.
24. Clark SL, Chan R, Zhao M, Xie LY, Copeland WE, Aberg KA, et al. Methylomic investigation of problematic adolescent cannabis use and its negative mental health consequences. J Am Acad Child Adolesc Psychiatry. 2021;60:1524–32.
25. van Dongen J, Bonder MJ, Dekkers KF, Nivard MG, van Iterson M, Willemsen G, et al. DNA methylation signatures of educational attainment. NPJ Sci Learn. 2018;3:7.
26. Karlsson Linner R, Marioni RE, Rietveld CA, Simpkin AJ, Davies NM, Watanabe K, et al. An epigenome-wide association study meta-analysis of educational attainment. Mol Psychiatry. 2017;22:1680–90.
27. van Dongen J, Hagenbeek FA, Suderman M, Roetman PJ, Sugden K, Chiocchetti AG, et al. DNA methylation signatures of aggression and closely related constructs: a meta-analysis of epigenome-wide studies across the lifespan. Mol Psychiatry. 2021;26:2148–62.
28. Karabegovic I, Portilla-Fernandez E, Li Y, Ma J, Maas SCE, Sun D, et al. Epigenome-wide association meta-analysis of DNA methylation with coffee and tea consumption. Nat Commun. 2021;12:2830.
29. Sandler DP, Hodgson ME, Deming-Halverson SL, Juras PS, D'Aloisio AA, Suarez LM, et al. The sister study cohort: baseline methods and participant characteristics. Environ Health Perspect. 2017;125:127003.
30. Stewart PA, Stenzel MR, Ramachandran G, Banerjee S, Huynh TB, Groth CP, et al. Development of a total hydrocarbon ordinal job-exposure matrix for workers responding to the Deepwater Horizon disaster: The GuLF STUDY. J Expo Sci Environ Epidemiol. 2018;28:223–30.
31. Boomsma DI, de Geus EJ, Vink JM, Stubbe JH, Distel MA, Hottenga JJ, et al. Netherlands Twin Register: from twins to twin families. Twin Res Hum Genet. 2006;9:849–57.
32. Justice AC, Dombrowski E, Conigliaro J, Fultz SL, Gibson D, Madenwald T, et al. Veterans Aging Cohort Study (VACS): overview and description. Med Care. 2006;44:S13–24.
33. Kaprio J. The Finnish Twin Cohort Study: an update. Twin Res Hum Genet. 2013;16:157–62.
34. Relton CL, Gaunt T, McArdle W, Ho K, Duggirala A, Shihab H, et al. Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). Int J Epidemiol. 2015;44:1181–90.
35. Moayyeri A, Hammond CJ, Hart DJ, Spector TD. The UK Adult Twin Registry (TwinsUK Resource). Twin Res Hum Genet. 2013;16:144–9.
36. Conway KP, Vullo GC, Kennedy AP, Finger MS, Agrawal A, Bjork JM, et al. Data compatibility in the addiction sciences: an examination of measure commonality. Drug Alcohol Depend. 2014;141:153–8.
37. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, et al. Epigenetic signatures of cigarette smoking. Circ Cardiovasc Genet. 2016;9:436–47.

38. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genome-wide association scans. Bioinformatics. 2010;26:2190–1.

39. Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, et al. rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. Genomics Proteom Bioinforma. 2021;19:619–28.

40. Odintsova VV, Suderman M, Hagenbeek FA, Caramaschi D, Hottenga JJ, Pool R, et al. DNA methylation in peripheral tissues and left-handedness. Sci Rep. 2022;12:5606.

41. Lee SH, Goddard ME, Wray NR, Visscher PM. A better coefficient of determination for genetic profile analysis. Genet Epidemiol. 2012;36:214–24.

42. Li M, Zou D, Li Z, Gao R, Sang J, Zhang Y, et al. EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. Nucleic Acids Res. 2019;47:D983–D988.

43. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. Epigenetics. 2015;10:1024–32.

44. Min JL, Hemani G, Hannon E, Dekkers KF, Castillo-Fernandez J, Luijk R, et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. Nat Genet. 2021;53:1311–21.

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

46. Xu Z, Xie C, Taylor JA, Niu L. ipDMR: identification of differentially methylated regions with interval P-values. Bioinformatics. 2021;37:711–3.

47. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8:1826.

48. Christiansen C, Castillo-Fernandez JE, Domingo-Relloso A, Zhao W, El-Sayed Moustafa JS, Tsai PC, et al. Novel DNA methylation signatures of tobacco smoking with trans-ethnic effects. Clin Epigenetics. 2021;13:36.

49. Markunas CA, Hancock DB, Xu Z, Quach BC, Fang F, Sandler DP, et al. Epigenome-wide analysis uncovers a blood-based DNA methylation biomarker of lifetime cannabis use. Am J Med Genet B Neuropsychiatr Genet. 2020;186:173–82.

50. Nannini DR, Zheng Y, Joyce BT, Kim K, Gao T, Wang J, et al. Genome-wide DNA methylation association study of recent and cumulative marijuana use in middle aged adults. Mol Psychiatry. 2023;28:2572–82.

51. Lohoff FW, Clarke TK, Kaminsky ZA, Walker RM, Bermingham ML, Jung J, et al. Epigenome-wide association study of alcohol consumption in N = 8161 individuals and relevance to alcohol use disorder pathophysiology: identification of the cystine/glutamate transporter SLC7A11 as a top target. Mol Psychiatry. 2022;27:1754–64.

52. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017;541:81–6.

53. Zhang J, Pan T, Zhou W, Zhang Y, Xu G, Xu Q, et al. Long noncoding RNA LINC01132 enhances immunosuppression and therapy resistance via NRF1/DPP4 axis in hepatocellular carcinoma. J Exp Clin Cancer Res. 2022;41:270.

54. Kunishima S, Okuno Y, Yoshida K, Shiraishi Y, Sanada M, Muramatsu H, et al. ACTN1 mutations cause congenital macrothrombocytopenia. Am J Hum Genet. 2013;92:431–8.

55. Li Y, Shu J, Cheng Y, Zhou X, Huang T. Identification of key biomarkers in Angelman syndrome by a multi-cohort analysis. Front Med. 2022;9:963883.

56. Xie T, Dong J, Zhou X, Tang D, Li D, Chen J, et al. Proteomics analysis of lysine crotonylation and 2-hydroxyisobutyrylation reveals significant features of systemic lupus erythematosus. Clin Rheumatol. 2022;41:3851–8.

57. Glessner JT, Chang X, Mentch F, Qu H, Abrams DJ, Thomas A, et al. COVID-19 in pediatrics: genetic susceptibility. Front Genet. 2022;13:928466.

58. Voisin S, Almen MS, Zheleznyakova GY, Lundberg L, Zarei S, Castillo S, et al. Many obesity-associated SNPs strongly associate with DNA methylation changes at proximal promoters and enhancers. Genome Med. 2015;7:103.

59. Ferreira MAR, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Eleven loci with new reproducible genetic associations with allergic disease risk. J Allergy Clin Immunol. 2019;143:691–9.

60. Campos AI, Kho P, Vazquez-Prada KX, Garcia-Marin LM, Martin NG, Cuellar-Partida G, et al. Genetic susceptibility to pneumonia: a GWAS meta-analysis between the UK Biobank and FinnGen. Twin Res Hum Genet. 2021;24:145–54.

61. Hemissi I, Boussetta S, Dallali H, Hellal F, Durand G, Voegele C, et al. Development of a custom next-generation sequencing panel for the determination of bladder cancer risk in a Tunisian cohort. Mol Biol Rep. 2022;49:1233–58.

62. Lee JW, Huang BX, Kwon H, Rashid MA, Kharebava G, Desai A, et al. Orphan GPR110 (ADGRF1) targeted by N-docosahexaenoylethanolamine in development of neurons and cognitive function. Nat Commun. 2016;7:13123.

63. Kveiborg M, Albrechtsen R, Couchman JR, Wewer UM. Cellular roles of ADAM12 in health and disease. Int J Biochem Cell Biol. 2008;40:1685–702.

64. Meng C, Zhou JQ, Liao YS. Autophagy-related long non-coding RNA signature for ovarian cancer. J Int Med Res. 2020;48:300060520970761.

65. Zhu W, Xiao X, Chen J. Silencing of the long noncoding RNA LINC01132 alleviates the oncogenicity of epithelial ovarian cancer by regulating the microRNA4315p/SOX9 axis. Int J Mol Med. 2021;48:1–12.

66. ElTelbany A, Khoudari G, Al-Khadra Y, McCullough A, Alkhouri N. Lower rates of hepatocellular carcinoma observed among cannabis users: a population-based study. Cureus. 2022;14:e24576.

67. Koltai H, Shalev N. Anti-cancer activity of cannabis sativa phytocannabinoids: molecular mechanisms and potential in the fight against ovarian cancer and stem cells. Cancers. 2022;14:4299.

68. Barrie AM, Gushue AC, Eskander RN. Dramatic response to Laetrile and cannabidiol (CBD) oil in a patient with metastatic low grade serous ovarian carcinoma. Gynecol Oncol Rep. 2019;29:10–12.

69. Biao T, Cai-Feng H, Xiao-Hong L, Xiao-Li C, Wen-Bei L, Jun W, et al. From Bowen disease to cutaneous squamous cell carcinoma: eight markers were verified from transcriptomic and proteomic analyses. J Transl Med. 2022;20:416.

70. Zhang L, Li X, Wan C, Da W, Zhang J, Fan L, et al. Bioinformatics analysis identification of AKT3 and RAC1 as key genes in postmenopausal osteoporosis. Exp Ther Med. 2022;24:656.

71. Gu B, Zhu M, Glass MR, Rougie M, Nikolova VD, Moy SS, et al. Cannabidiol attenuates seizures and EEG abnormalities in Angelman syndrome model mice. J Clin Invest. 2019;129:5462–7.

72. Zavala-Tecuapetla C, Luna-Munguia H, Lopez-Meraz ML, Cuellar-Herrera M. Advances and challenges of cannabidiol as an anti-seizure strategy: preclinical evidence. Int J Mol Sci. 2022;23:16181.

73. Clouse G, Penman S, Hadjiargyrou M, Komatsu DE, Thanos PK. Examining the role of cannabinoids on osteoporosis: a review. Arch Osteoporos. 2022;17:146.

74. Bachari A, Piva TJ, Salami SA, Jamshidi N, Mantri N. Roles of cannabinoids in melanoma: evidence from in vivo studies. Int J Mol Sci. 2020;21:6040.

75. Rodriguez Mesa XM, Moreno Vergara AF, Contreras Bolanos LA, Guevara Moriones N, Mejia Pineros AL, Santander Gonzalez SP. Therapeutic prospects of cannabinoids in the immunomodulation of prevalent autoimmune diseases. Cannabis Cannabinoid Res. 2021;6:196–210.

76. Liu W, Xu J, Dai T, Shen G, Feng J. An integrating strategy for serum metabolomics and microarray analysis to expand the understanding of diet-induced obesity. Anal Methods. 2021;13:3127–35.

77. Clark TM, Jones JM, Hall AG, Tabner SA, Kmiec RL. Theoretical explanation for reduced body mass index and obesity rates in cannabis users. Cannabis Cannabinoid Res. 2018;3:259–71.

78. Yayan J, Rasche K. Damaging effects of cannabis use on the lungs. Adv Exp Med Biol. 2016;952:31–34.

79. Park T, Chen H, Kim HY. GPR110 (ADGRF1) mediates anti-inflammatory effects of N-docosahexaenoylethanolamine. J Neuroinflammation. 2019;16:225.

80. Abdulkareem NM, Bhat R, Qin L, Vasaikar S, Gopinathan A, Mitchell T, et al. A novel role of ADGRF1 (GPR110) in promoting cellular quiescence and chemoresistance in human epidermal growth factor receptor 2-positive breast cancer. FASEB J. 2021;35:e21719.

81. Wang H, Wu H, Chen Y. Serum ADAM metallopeptidase domain 12 as a promising diagnostic biomarker for breast and liver cancer. Clin Transl Oncol. 2022;25:473–81.

82. Chung MW, Park YL, Park SY, Joo YE. Expression of ADAM12 in gastric cancer and its relation to tumor cell behavior and prognosis. Vivo. 2022;36:2194–204.

83. Park YL, Park SY, Oh HH, Chung MW, Hong JY, Kim KH, et al. A disintegrin and metalloprotease 12 promotes tumor progression by inhibiting apoptosis in human colorectal cancer. Cancers. 2021;13:1927.

84. Shao S, Li Z, Gao W, Yu G, Liu D, Pan F. ADAM-12 as a diagnostic marker for the proliferation, migration and invasion in patients with small cell lung cancer. PLoS ONE. 2014;9:e85936.

85. Frohlich C, Albrechtsen R, Dyrskjot L, Rudkjaer L, Orntoft TF, Wewer UM. Molecular profiling of ADAM12 in human bladder cancer. Clin Cancer Res. 2006;12:7359–68.

86. McCartney DL, Hillary RF, Stevenson AJ, Ritchie SJ, Walker RM, Zhang Q, et al. Epigenetic prediction of complex traits and death. Genome Biol. 2018;19:136.

87. Kopjar N, Fuchs N, Zunec S, Mikolic A, Micek V, Kozina G et al. DNA damaging effects, oxidative stress responses and cholinesterase activity in blood and brain of wistar rats exposed to delta(9)-tetrahydrocannabinol. Molecules. 2019;24.

88. Wolff V, Schlagowski AI, Rouyer O, Charles AL, Singh F, Auger C, et al. Tetra-hydrocannabinol induces brain mitochondrial respiratory chain dysfunction and increases oxidative stress: a potential mechanism involved in cannabis-related stroke. Biomed Res Int. 2015;2015:323706.

89. Roland AB, Ricobaraza A, Carrel D, Jordan BM, Rico F, Simon A, et al. Cannabinoid-induced actomyosin contractility shapes neuronal morphology and growth. eLife. 2014;3:e03159.

90. Saunders GRB, Wang X, Chen F, Jang SK, Liu M, Wang C, et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. Nature. 2022;612:720–4.

91. Michels KB, Binder AM, Dedeurwaerder S, Epstein CB, Greally JM, Gut I, et al. Recommendations for the design and analysis of epigenome-wide association studies. Nat Methods. 2013;10:949–55.

92. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinforma. 2012;13:86.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

All authors reviewed the results and approved the manuscript. FF, BQ, DBH, and EOJ conceived and designed the study. FF, BQ, KGL, JvD, JAM, SL, ML, VVO, RC, ZX, LZ, MM, and YX conducted cohort-specific analyses. FF, JAM, and LZ performed meta-analysis. MO, JAT, JTB, JK, DIB, KX, and DPS oversaw the cohort research team and contributed to study design, analysis plan and interpretation of the results. JMV and LJB provided consulting to study design, analysis plan and interpretation of the results.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41380-023-02310-w.

Correspondence and requests for materials should be addressed to Fang Fang.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.