Genetics in Medicine (2025) 27, 101586

ELSEVIER

Genetics in Medicine
An Official Journal of the ACMG

www.journals.elsevier.com/genetics-in-medicine

ARTICLE

Discovery of a DNA methylation episignature as a molecular biomarker for fetal alcohol syndrome

Liselot van der Laan[1,2], Raissa Relator[3], Irene Valenzuela[4], Adri N. Mul[1,2], Mariëlle Alders[1,2], Michael A. Levy[3], Jennifer Kerkhof[3], Jessica Rzasa[3], Anna M. Cueto-González[4], Amaia Lasa-Aranzasti[4], Cristina Cea-Arestin[4], Marcel M.A.M. Mannens[1,2], Mieke M. van Haelst[1,2,5,6], Eduardo F. Tizzano[4], Bekim Sadikovic[3,7,*], Peter Henneman[1,2,*]

ARTICLE INFO

Article history:
Received 7 December 2024
Received in revised form
30 August 2025
Accepted 12 September 2025
Available online 18 September 2025

Keywords:
FASD
DNA methylation
Episignature
Fetal alcohol spectrum disorder
Molecular biomarker

ABSTRACT

Purpose: Fetal alcohol spectrum disorder (FASD) encompasses a range of clinical features and neurodevelopmental disorders in children exposed to alcohol in utero. Despite its global public health significance, FASD diagnosis remains challenging because of nonspecific clinical findings and the lack of an accurate molecular diagnostic biomarker. This study aimed to evaluate peripheral blood DNA methylation (DNAm) profiles as a potential diagnostic biomarker for fetal alcohol syndrome.

Methods: Genomic DNAm profiles from 93 individuals with suspected or confirmed FAS, including a clinically diagnosed FAS subgroup, were analyzed and compared with a large database of control and patient cohorts with previously reported DNAm episignatures. Functional analysis of these DNAm profiles was performed to identify episignatures and assess their potential diagnostic utility.

Results: A relatively sensitive and specific DNAm episignature for FAS was identified. Comparative epigenomic analysis revealed functional correlations between FAS and other rare genetic disorders, supporting the robustness of the identified DNAm profiles as a diagnostic tool.

Conclusion: This study demonstrates that unique DNAm profiles provide a robust episignature biomarker for FAS. These findings contribute to the molecular understanding of FAS and hold promise for improving diagnostic accuracy for this complex disorder.

The Article Publishing Charge (APC) for this article was paid by Peter Henneman.

Liselot van der Laan, Raissa Relator, and Irene Valenzuela jointly coordinated this work.

Eduardo F. Tizzano, Bekim Sadikovic, and Peter Henneman jointly coordinated this work.

*Correspondence and requests for materials should be addressed to Bekim Sadikovic, Department of Pathology and Laboratory Medicine, Western University, 1151 Richmond St, London, Ontario N6A 3K7, Canada. *Email address:* bekim.sadikovic@lhsc.on.ca OR Peter Henneman, Department of Human Genetics, Amsterdam 1105AZ, Amsterdam, The Netherlands. *Email address:* p.henneman@amsterdamumc.nl

Affiliations are at the end of the document.

doi: https://doi.org/10.1016/j.gim.2025.101586

Introduction

Fetal alcohol spectrum disorder (FASD) is a neurodevelopmental disorder caused by prenatal alcohol exposure, characterized by developmental and growth delays, including microcephaly and growth retardation. Moreover, FASD involves central nervous system impairments such as seizures, cognitive, and behavioral deficits, memory and learning difficulties, hyperactivity, impulsivity, and distinct facial abnormalities. These facial abnormalities include small palpebral fissures, a thin upper lip, a smooth philtrum and low-set ears.[1,2] The neurobehavioral phenotype of FASD can be debilitating and significantly affect daily life, affecting cognitive function, social interactions, and academic performance. The accepted standard for FASD diagnosis is the 4-digit diagnostic code, which is based on the scoring of 4 key diagnostic features: (1) prenatal alcohol exposure, (2) growth failure, (3) FAS-specific facial abnormalities, and (4) neurological features, such as cognitive and motor delays.[3] Using these diagnostic criteria, patients may be diagnosed with fetal alcohol syndrome,[4] partial FAS (pFAS), alcohol-related neurodevelopmental disorder,[5] or alcohol-related birth defects.[5-7] Fetal alcohol syndrome[4] represents the most severe affected patients.

FASD diagnosis remains difficult because of both its heterogeneous presentation and clinical overlap with many genetic conditions. Symptoms such as cognitive impairments, behavioral issues, and growth abnormalities are not unique to FASD, making it difficult to distinguish from other neurodevelopmental or genetic disorders.[8] As a result, many individuals undergo extensive genetic testing, including microarrays and exome sequencing.[8-10] The lack of symptom penetrance and variable expressivity in FASD further complicates its diagnosis because not all individuals may present the full spectrum of symptoms.[11] As a result, FASD often becomes a diagnosis of exclusion, reached only after other potential genetic causes have been ruled out.[12]

FASD is a significant public health issue, with prevalence rates ranging between 1 and 5%, which is much higher than any other teratogenic disorder. However, these rates vary significantly between countries, with some regions experiencing particularly high prevalence, such as South Africa, Ireland, Italy, Croatia, Russia, and Poland.[13-15] Furthermore, significantly higher rates of FASD have been observed among children in foster care or correctional facilities.[16] The development of molecular biomarkers could reduce reliance on lengthy diagnostic processes and improve early identification.[17] Without reliable biomarkers, many individuals with FASD may remain undiagnosed or misdiagnosed, leading to delays in receiving appropriate interventions and support, such as behavioral therapies, individualized education support, and medical care for cooccurring conditions. Moreover, biomarkers-based tools can support accurate FAS diagnosis, leading to tailored family training that improves long-term outcomes.

Accurate diagnosis is crucial for ensuring access to timely interventions and support services, and for informing families of future pregnancy risks.[7] Because FASD is caused by prenatal alcohol exposure rather than genetic factors, the risk of recurrence in future pregnancies is low, as long as alcohol is avoided during pregnancy.[18,19] However, overcoming alcohol addiction can be extremely challenging, with a high risk of relapse. This makes ongoing support and intervention critical to preventing future cases of FASD and ensuring the health of both the mother and future children. Understanding these dynamics can help families and health care providers better manage and mitigate risks.[20]

Establishing an FASD diagnosis is challenging for the following reasons: (1) ambiguous clinical features that overlap with other rare disorders, (2) lack of accurate records of prenatal alcohol exposure, and (3) lack of a molecular biomarker.[21] Furthermore, many individuals with FASD lack a confirmed history of prenatal alcohol exposure, making diagnosis reliant on clinical presentation alone. This highlights the urgent need for objective molecular biomarkers to support FASD diagnosis.

Epigenetics refers to the study of heritable modifications to the genome that do not involve alterations to the DNA sequence. DNA methylation (DNAm) is the most extensively studied epigenetic mechanism, and it is known that many rare genetic disorders are associated with distinct DNAm patterns known as episignatures.[22] In recent years, a growing number of episignatures have been developed as stable and reliable biomarkers for the diagnosis of congenital genetic disorders and for the reclassification of variants of uncertain significance.[23-28] These episignatures have been successfully implemented in clinical diagnostic laboratories, offering significant diagnostic utility in genetically unresolved patients with suspected rare disorders.[29]

Diagnostic episignature biomarkers have been developed for more than 100 genetic disorders,[22] many of which exhibit significant phenotypic overlap. Additionally, episignatures have recently been discovered in syndromic disorders with no known genetic etiology, such as VACTERL (vertebral defects, anal atresia, cardiac defects, trachea-esophageal fistula, renal anomalies, and limb abnormalities) association and oculoauriculovertebral spectrum association.[30] A robust episignature was also reported in the fetal valproate syndrome representing the first major evidence of a DNAm biomarker associated with a prenatal exposure to a teratogen, in this case, an antiepileptic medication valproate.[30,31]

For FASD, significant DNAm aberrations have been previously reported in several genome-wide epigenetic association studies.[1,32-34] Consequently, we hypothesized that individuals with FASD may display a robust DNAm episignature biomarker as well. Developing an episignature for FASD could provide an objective molecular tool to support clinical diagnosis and improve patient outcomes.

In this study, we aimed to identify a sensitive and specific DNAm episignature in individuals within the spectrum of fetal alcohol disorders that can be incorporated into diagnostic testing.

Materials and methods

Study cohort

All FASD diagnoses and 4-digit scoring were performed by at least 2 clinical geneticists within the Amsterdam University Medical Center, The Netherlands; Medical University of Wroclaw, Poland; and the University Hospital Vall d'Hebron, Barcelona, Spain. Cross-institute confirmation of FAS diagnosis ($N = 28$) by comparing the diagnostic criteria was performed for each case within the training set. Informed consent was obtained from all participants and/or legal guardians. The study was approved by the Medical Ethical Committees of aforementioned institutes. The analysis cohort includes 93 individuals with suspected FASD diagnosis, including 27 females and 66 males with ages ranging from 4.8 months to 37 years (median = 10). Four-digit diagnostic scores[6,7] for the 4 key diagnostic features of FASD (GROWTH = growth deficiency, FACE = facial characteristics, CNS = brain damage/dysfunction, and ALC = gestational alcohol exposure) were obtained for 87 individuals, whereas the remaining 6 had no recorded scores.

Sample processing

All samples analyzed in this study were anonymized, and DNA was extracted from peripheral blood samples. DNA concentration was measured using Qbit fluorometric quantification, according to manufacturer's (ThermoFisher Scientific) protocol. DNAm profiles were generated by bisulfite-conversion of 500 ng DNA and subsequent analyses using Illumina Infinium HumanMethylation450K, Illumina Infinium MethylationEPIC v1 and EPICv2, BeadChip arrays following manufacturer's protocol (Illumina). Before data analyses quality control was performed as follows: raw intensity data files containing methylated and unmethylated signal intensities were uploaded in R (version 4.1.2) and normalized using the SeSAMe package,[35] with dye bias correction and background subtraction. Quality probe masking[36] and detection P value probe masking ($P = .05$) based on Infinium out-of-band signal calibration were also implemented as part of the standard preprocessing. Additional quality control measures were carried out including assessment of a methylation density plots and concordance between recorded and predicted age and sex. A total of 109 DNAm profiles were included in this study, comprising 89 HM450K array profiles and 20 EPICv1 array profiles. In addition, technical replicate DNAm profiles were generated for 16 samples using EPICv2. These were annotated as such to test the robustness of the signature across array types.

Signature probe selection and classifier development

The episignature training and analysis were performed following previously published protocols using the EpiSign Discovery Software.[22,37] Computations were done using the R (version 2.4.1) statistical software and associated libraries. Before signature detection analysis, the following probes were removed from the data set: Probes annotated on the X or Y chromosome, probes known to be cross-reactive, probes affected by SNV,[38,39] probes showing a high variability after manufacturing change from prior methylation array products by Illumina, and EPICv1 probes that were removed in the design of the EPICv2 array. The afore mentioned filtering steps resulted in exclusion of in total 169,708 unique probes. Next, differential methylation analysis was performed on matched cases and controls. Controls were selected from the EpiSign Knowledge Database (EKD) using a case-control ratio of 1:2. Selection was executed using the R MatchIt package (version 4.5.1)[40] by matching age, sex, batch, and array type information. Principal component analysis was implemented to examine data structure and to identify outliers before signature analysis. Methylation beta values were logit-transformed to M-values and subsequently fitted in a multivariate linear regression model, implemented using the R limma package (version 3.50.0),[41] with methylation levels as predictors, case/control labels as response, and estimated blood cell composition as covariates in the model. Moderated t-statistics and P values were computed using empirical Bayes, and false discoveries were minimized by adjusting statistics using the Benjamini-Hochberg procedure. The signature was defined by using the top probes resulting from a 3-staged selection. First, the probes were rank selected by decreasing value of the probe score, defined as the product of absolute mean methylation difference and the negative log of adjusted P values. Then, the probes with the highest variable importance, computed using the area under the receiver operating characteristic (AUROC) curve were selected. Finally, probes with high pairwise correlations were excluded from further analyses. Combinations of the probe score, AUROC, and pairwise correlation cutoffs were used to generate probe sets. Each probe set was investigated by unsupervised clustering of training cases and controls using Euclidean Clustering (heatmaps) and multidimensional scaling (MDS). Plots were assessed and the probe set producing optimal segregation in clustering plots were selected for the episignature. To confirm reproducibility and robustness of the approach, leave-one-out cross-validation was performed on the training cases and clustering concordance of each hold-out test case was evaluated. A binary classifier was developed using a support vector machine model implemented using the R e1071 package

(version 1.7-13). The selected signature probes were used as features and training cases and controls were used for training. Reference EKD samples not used in the training stage were also added as training controls to increase model specificity. Four-fold cross-validation was performed using a 75% to 25% train-test split of the nontraining EKD samples and support vector machine (SVM) scores of the test samples were averaged over the 4 sets to assess model performance. The hyperparameters for the SVM model, such as kernel function, regularization parameter and class weights were determined via grid search and nested cross-validation using the training data.

Functional annotation and comparison with other episignature-positive rare disease cohorts

Functional annotation, correlation, and comparison with other previously described EpiSign episignature cohorts were performed following previously published methods.[42] In brief, we matched the training cases to unaffected and EpiSign-negative controls, identified differentially methylated probes (DMPs), as well as regions (DMRs), and annotated results and investigated the similarity and overlap of the FAS signature with 56 episignature-positive disorder cohorts included in the EpiSign version 3 classifier. DMPs were identified using the same linear model for probe selection as in the training analysis, whereas DMR analysis was performed using the R DMRcate package (version 2.10.0).[43] Regions were defined to have at least 5 CpG sites in cases in which any 2 contiguous sites are within 1 kb of each other. DMRs were identified as regions with statistically significant (Fisher P value $< .01$) mean methylation difference of at least 5% between cases and controls. Differentially methylated sites and regions were annotated using CpG and gene-based annotations using annotatr (version 1.22.0)[44] and AnnotationHub (version 3.2.2) packages in R. Gene set enrichment analysis was performed on differential probes and regions using missMethyl package (version 1.30.0). Cohort similarities based on DMP overlaps were visualized using heatmap and scatter plot. Additionally, comparison of signatures using the top 500 DMPs ranked by P value for each cohort was illustrated using a tree and leaf plot in which the node sizes indicate the relative total number of DMPs and node colors reflect the global methylation profile of the signature. For signatures with less than 500 DMPs, all differentially methylated probes were used. Samples for each disorder cohort were aggregated and median value for each probe was calculated before performing hierarchical clustering using Ward's method and generating a tree and leaf plot.

Exome screening of ID genes in an FASD subset

To exclude the presence of any (likely) pathogenic genetic variant within genes known to be linked to intellectual deficit (ID), we performed exome sequencing (ES) in a subset of the cohort of 22 cases, of which 16 FAS were positive for the episignature, and 6 were negative. Libraries were prepared using the Kapa HTP kit (Illumina) and capture was performed using the KAPA HyperExome (Roche). Sequencing was done on an Illumina NovaSeq 6000 platform according to the manufacturer's protocols. Variant annotation and prioritizing were preformed following a targeted analysis approach of an ID gene panel (https://genoomdiagnostiek.nl/en/product-tag/wes-en/) (IDv14) were done using Alissa Interpret (Agilent), based on the human genome build GRCH37/hg19. Variants were considered that have an allele frequency $<$ 0.02% (dominant conditions) or $<$1% (recessive conditions), as well as variants listed as (likely) pathogenic in Human Gene Mutation Database or ClinVar databases or predicted to have a loss-of-function effect (nonsense, frameshift, start-loss, stop-loss, and canonical splice site variants).

Results

Defining an episignature for FAS

In this study, the following 5 FASD groups were defined (Table 1)[1]: a set of 28 participants with fetal alcohol syndrome (annotated as *FAS*), whose score for each key feature was at least 2 and showing a sum score ≥ 9.[45] A set of 16 participants with fetal alcohol syndrome (annotated as FAS?), whose score for each key features was at least 2 and showing a sum score ≥ 11. In contrast to group 1, these cases lack a confirmed clinical diagnosis of FAS and were classified as suspected based on available clinical features. This group shows a variable clinical presentation across all 4 key scores, unlike pFAS cases, for which exposure is typically lower, as shown in Table 1.[46] A selection of 24 participants with pFAS, indicated by a score in at least 1 key feature of 1 and for which we have not observed any known FASD comorbidity.[4] A group of 6 participants with prenatal alcohol exposure but without a clear FASD phenotype (annotated as no clear FASD pheno) Although four individuals had recorded 4-digit scores, the clinical evaluation of features was not considered completely consistent with FASD; therefore, we cannot exclude the possibility of involvement of other neurodevelopmental disorders with overlapping characteristics.[5] A set of 5 participants with suspected FASD and for which we have confirmed a pathogenic genetic variation associated with ID and/or for which we observed FASD nonspecific comorbidities (annotated as Withcomorbidities). Finally,[6] a set of 14 individuals with possible FASD (annotated as *FASD?* and *FASD?_v2*). Uncertainty in these cases was primarily due to incomplete or inconsistent clinical presentation, limited documentation of prenatal alcohol exposure, or borderline diagnostic scores that did not meet all criteria for definitive classification. Additional detailed description of the whole cohort and subsets is described in Supplemental Table 1.

Table 1 FASD cohort and subsets descriptions (please see more explanation in the text)

Cohort Annotation	Group	*N*	Sex M/F	Average Sum Digit Score	Group Description	Positive for FAS Episignature
FAS (training[a])	1	28	20/8	14	Confirmed FAS cases used for training the model	- (used for training)
FAS_rep (replicate EPICv2, testing)	1	16	10/6	14	Replicated FAS cases tested on EPICv2 array	16
FAS? (testing)	2	16	9/7	13	Cases with suspected but unconfirmed FAS, used for testing	6
pFAS (testing)	3	24	15/9	10	Partial FAS cases included in the testing group	8
no_clear_FASDpheno (testing)	4	6	6/0	12	Cases with prenatal alcohol exposure but without clear FASD phenotype	1
with comorbidities (testing)	5	5	4/1	12	Cases with FASD suspicion but also presenting additional comorbidities	2
FASD? (testing)	6	11	9/2	14	Cases with uncertain FASD diagnosis	5
FASD?_v2 (testing)	6	3	3/0	11	Additional uncertain FASD cases tested on EPICv2	3

F, female; *M*, male.
[a]Training set FAS, included in final episignatures training.

First, we performed signature training analysis based on only a set of 28 out of 44 samples with full fetal alcohol syndrome (group 1), for which cross-institutional clinical diagnosis was confirmed. We detected a common methylation profile as a result of iterative feature selection by first considering the top 900 probes ranked by *P* values and methylation difference. Subsequently, we selected the top 450 probes based on the AUROC score. Finally, we removed probes with pairwise correlation greater than 0.6. The final probe set for the DNAm signature included 204 probes with mean methylation differences ranging from 5% to 14% (median = 7%) between cases and controls. Fifty-two percent of these probes were hypermethylated (Supplemental Table 2). Next, unsupervised clustering of the training samples (Table 1) using the signature probes revealed clear separation of cases from controls in both heatmap and MDS plot (Figure 1A and B). As expected, we also observed high scores (>0.75) for the technical replicates of 16 training FAS cases, with evident clustering alongside training cases in both heatmap and MDS plots. Unsupervised leave-one-out cross-validation on the training cases confirmed the reproducibility and robustness of the identified signature as each hold-out test case groups with the remaining training cases for each cross-validation set (Supplemental Figure 1).

Development of the SVM classifier for FASD

We developed an SVM-based predictive model and assessed its utility in identifying individuals with FAS or with the overlapping phenotype spectrum. Signature probes were used as features. Training samples with 75% of non-training EKD samples (controls and other disorders) were used as training data, whereas the other 25% of nontraining EKD were used as testing. Methylation variant pathogenicity (MVP) scores of test samples were computed and

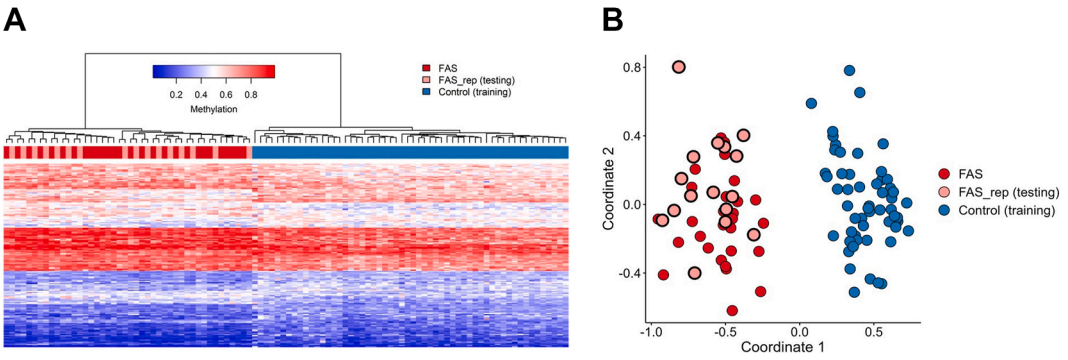

Figure 1 Discovery of the fetal alcohol spectrum disorder (FASD) episignature, based on group 1. A. The Euclidean hierarchical clustering heatmap shows each column as an FASD discovery case (highlighted in red). Each row represents a specific probe chosen for this episignature. A clear distinction is evident between the cases (in red) and EKD controls (in blue). B. The multidimensional scaling (MDS) plot displays the separation between FASD cases and controls. The pink replicate samples behave the same in both (A) and (B), clustering with the training cases and away from controls.

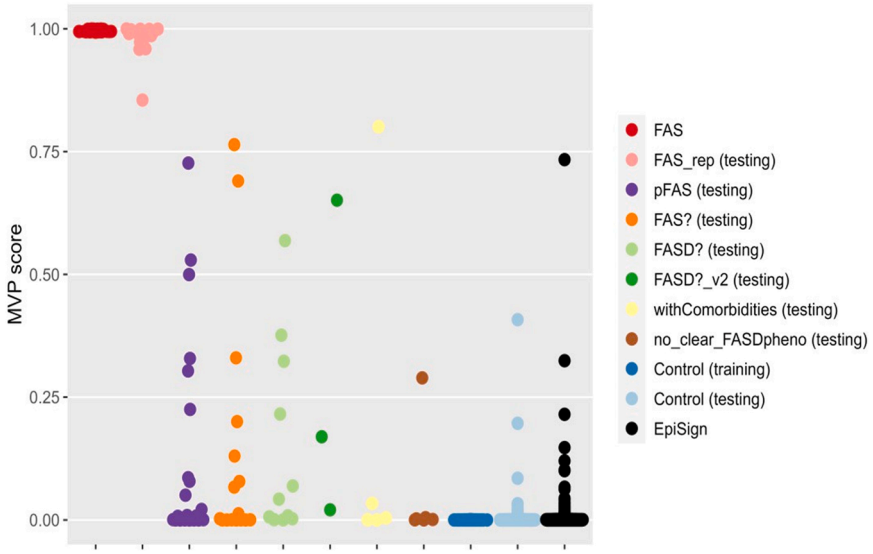

Figure 2 Methylation variant pathogenicity scores were also computed using the support vector machine (SVM)-based classifier and averaged over the 4-fold cross-validation. Control (testing) are non-NDD controls in EKD that were not used for signature training, whereas EpiSign label are for nontraining samples from the EKD that are positive for at least 1 signature.

averaged over the 4-fold cross-validation, in which a higher score suggests a more similar methylation profile to the training cases. The model showed specificity for FAS with vast majority of the EKD test controls obtaining scores below 0.3 (Figure 2, dots in light blue). In addition, other EpiSign disorders (annotated in black) generally showed similar low scores below, with the exception of one particular case.

We analyzed MVP scores and clustering behavior for each FASD subgroup as defined in Table 1.

Group 2 (FAS?): the 16 additional FAS? test cases (Figure 3A and B; dots in orange), showed that 6 clustered with confirmed FAS cases in the heatmap and MDS plots, although only 3 displayed MVP scores above 0.3.

Group 3 (pFAS): in the 24 partial FAS (pFAS) cases (Figure 3C and D; dots in purple), 8 clustered with the FAS group in both heatmap and MDS plot. These 8 showed MVP scores between 0.3 and 0.7, indicating a mild DNAm overlap with the FAS episignature.

Group 4 (no clear FASD phenotype): for the 6 participants with prenatal alcohol exposure but without a clearly defined FASD phenotype (Figure 3E and F; dots in brown), we found that only 1 individual clustered with FAS cases and exhibited a high MVP score of 0.8.

Group 5 (with comorbidities): among the 5 participants with suspected FASD and known comorbidities (Figure 3E and F; dots in yellow), 2 showed partial clustering with FAS cases in heatmap and MDS plots, but none had MVP scores above 0.3.

Group 6 (FASD? and FASD?_v2): in this group of 14 participants with uncertain FASD diagnosis (Figure 3G and H; dots in green), 8 clustered with FAS cases in both heatmap and MDS plots. However, only 3 had MVP scores exceeding 0.3.

Detailed information regarding the MVP scores of all tested subsets and participants is described Supplemental Table 1.

Next, using the χ^2 test, we investigated if any of the key features, ie, prenatal alcohol exposure, growth retardation, FAS-specific facial abnormalities, and neurological features were correlated with the MVP scores or their prediction class. For this analysis, we focused on nonreplicate samples from 2 subsets: confirmed FAS cases (discovery samples) and pFAS investigational samples, totaling 40 individuals. Our analysis indicated that fetal alcohol exposure (ALC) scores might influence the prediction in the pFAS group ($P = .04$). Additionally, when analyzing the combined cohort, we observed a possible association between prediction class and growth phenotype score ($P = .03$). This trend was particularly evident in samples with higher scores for these features, which had higher MVP prediction scores (Supplemental Figure 2). However, further analysis using the Spearman correlation test between MVP scores and key diagnostic features scores did not reach statistical significance ($P > .05$).

Genome-wide methylome annotation and comparison with EKD cohorts

We performed feature annotation analyses of the FAS signature cohort following existing protocols published by our laboratory.[42] A global differentially methylated profile was identified by comparing training cases with matched samples selected from the EKD. The matched samples included healthy controls and unresolved cases that tested negative for any of the 56 episignature disorders included in the EpiSign v3 classifier. We identified a total of 1894

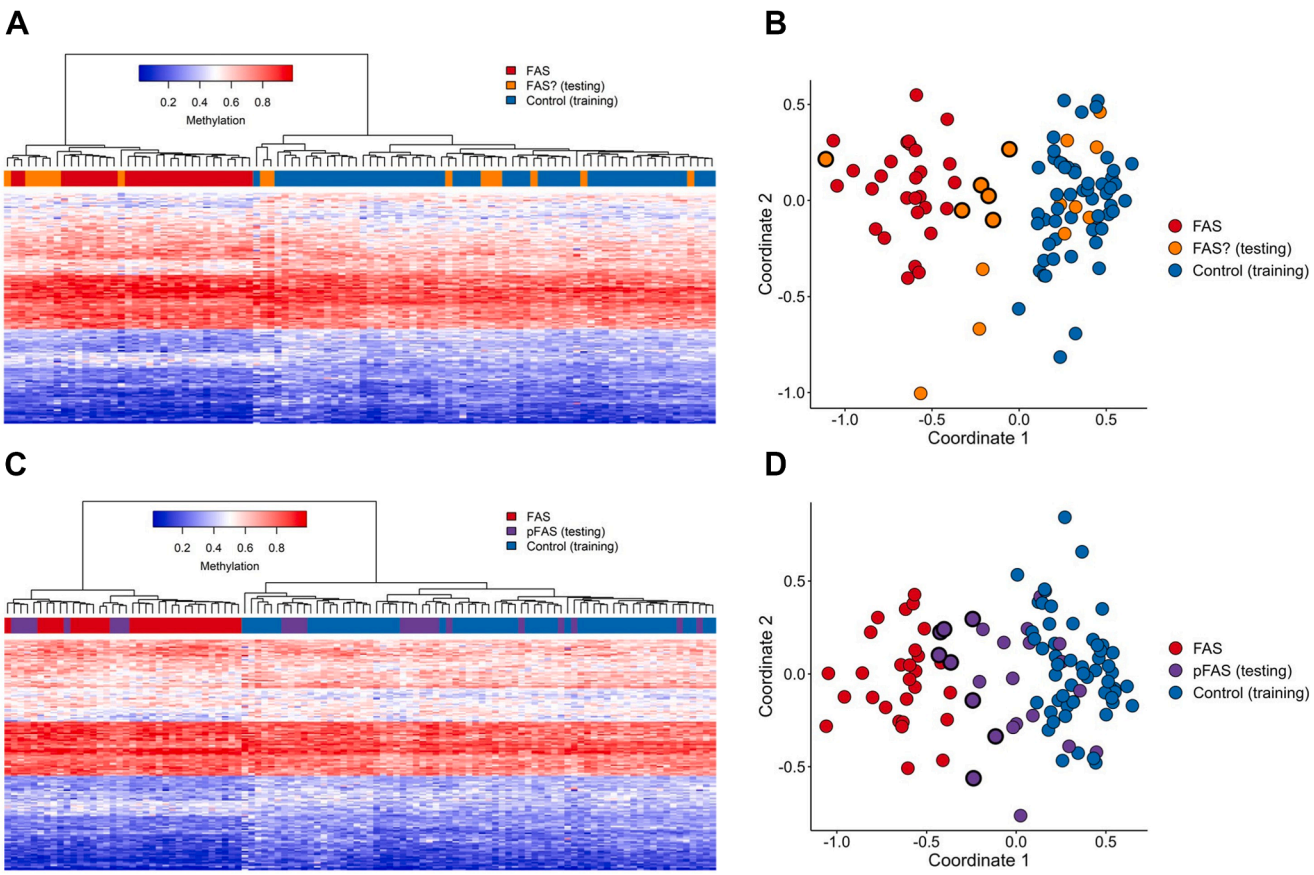

Figure 3 Episignature assessment of cases with suspected fetal alcohol spectrum disorder. Testing cohort samples were evaluated using the FAS episignature model. Unsupervised hierarchical clustering and multidimensional scaling analysis of possible FAS samples (A and B; orange, group 2), partial FAS samples (C and D; purple, group 3), suspected FASD samples but found no clear FASD features (E and F; brown, group 4), suspected FASD samples but with confirmed comorbidities (E and F; yellow, group 5), and possible FASD samples (G- and H; green, group 6) were performed to evaluate whether test samples will cluster with training FAS cases (red) or matched controls. Boldly outlined dots indicate samples classified as positive for the FAS episignature.

differentially methylated probes with a mean methylation difference ranging between 5% and 14%. Moreover, this set of DMPs showed predominantly hypermethylation (64% DMPs) in FAS cases (Supplemental Figure 3A, Supplemental Table 3). Genomic location annotation of DMPs revealed that the highest percentage of probes selected are in CpG islands (annotation distribution: island = 34%, inter-CGI = 33%, shore = 27%, shelf = 6%) and in coding sequence regions (CDS) in relation to genes (annotation distribution: CDS = 37%, intergenic = 26%, promoter = 26%, promoter+ = 11%) (Supplemental Figure 3B and C). These annotation distributions were found to be statistically different from the baseline distributions calculated using all filtered probes before analysis, with $P < .0001$ using the χ^2 goodness of fit test. Gene set enrichment analysis of these DMPs compared with the universal probe set preanalysis showed no significant gene ontology and human phenotype ontology (data not shown). However, our Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis found a significant number of DMPs overlap with genes associated to the cAMP signaling pathway (ID hsa04024; FDR $P = .0237$, 32 genes overlapped).

Further inspection of signature correlations showed significant overlapping DMPs between FAS and EpiSign v3 cohorts, with the highest percentage of FAS DMPs being shared with Sotos (*NSD1*, 49%), ICF1 (*DNMT3B*, 41%) and Dup7 (chr7q11.23dup, 38%) (Figure 4A). A tree and leaf plot illustrating similarity of disorder cohorts based on the top 500 differentially methylated probe for each signature group was generated by clustering cohorts using the median values for the aggregated top DMPs. Our results show that among the 56 cohorts, FAS is most similar to BAFopathy (includes *ARID1A*, *ARIA2B*, *SMARCA2*, *SMARCA4*, and *SMARCB1*), MRD23 (*SETD5*), and Kabuki (*KMT2D* and *KDM6A*) (Figure 4B).

Differentially methylated regions

In total, we have identified 24 DMRs with at least 5% mean methylation difference between FAS cases and unaffected controls (Supplemental Table 4). Two of the 24 DMRs were hypomethylation events, whereas the remaining 22 DMRs were hypermethylated. CpG sites included in these regions were enriched in CpG islands (annotation distribution:

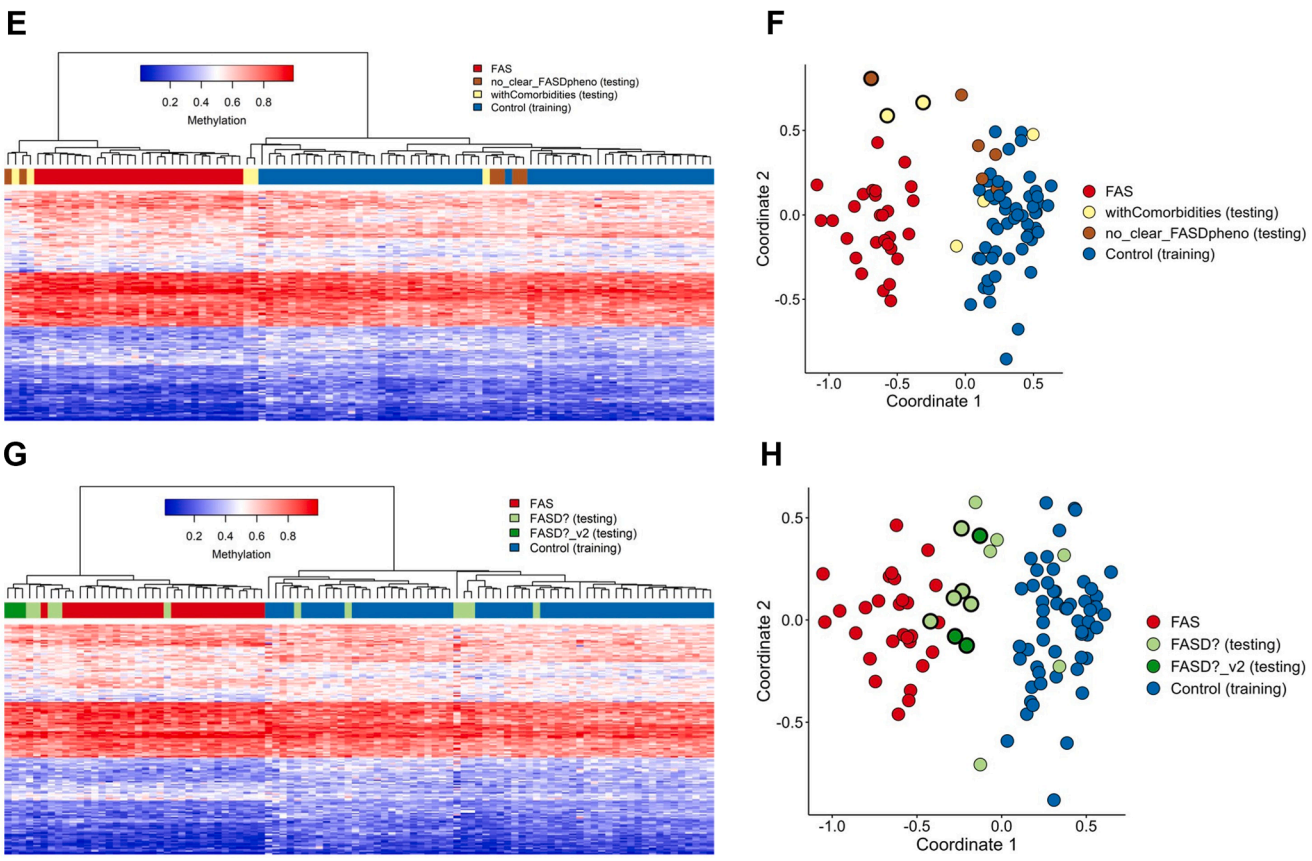

Figure 3 Continued.

island = 67%, inter-CGI = 21%, shelf = 8%, shore = 4%) and promoter regions (annotation distribution: promoter = 37%, CDS = 33%, intergenic = 17%, promoter+ = 13%) (Supplemental Figure 3B and C). We did not find any significant Gene Ontology, Human Phenotype Ontology, or pathway terms enriched in the DMR results (data not shown).

Exome screening of ID genes in an FASD subset

To assess the presence of any pathogenic gene variant linked to ID, we performed ES on a subset of the cohort. This subset included 22 cases, with 16 FAS cases having a positive signature and 6 participants having a negative signature. In this subset, we found 1 likely pathogenic variant in 1 case; case 39 (FAS_rep, group 1), who harbored a *BRPF1* variant. This case was positive for the FAS episignature. Notably, this case also corresponds to case 12 from the original training cohort (group 1) used to develop the FAS episignature. Although the inclusion of this individual is not expected to have biased the signature given its robust performance in leave-one-out-cross-validation we report it here for clarity. The remaining 21 cases, of which 15 showed positive signatures, did not show any (likely) pathogenic variants in genes linked to ID. Five cases had a negative signature and no variants in known ID genes (Supplemental Table 5).

Discussion

We identified a relatively sensitive and specific DNAm episignature for individuals with FAS, by generating DNAm profiles from peripheral blood of 93 clinically characterized or suspected individuals with FASD. The episignature was developed using data from 28 clinically confirmed FAS participants and subsequently tested on different FASD subsets, including a technical replicate set of 16 individuals and participants with suspected FAS. Cross-validation demonstrated that all FAS cases clustered together, indicating that the FAS DNAm episignature is robust and reproducible. The SVM model confirmed the selected probes as strong biomarkers for FAS, showing high specificity and sensitivity relative to other episignature-positive rare neurodevelopmental disorders. In testing with 65 participants (groups 2-6) exhibiting various FASD phenotypes, we classified 25 participants as positive for the FAS episignature based on the MVP score threshold established through our classification model. Regarding the performance of the different FASD subgroups tested, we found that although some groups (such as pFAS and FAS?) exhibited partial overlap with the FAS episignature, the level of concordance varied. For instance, a number of individuals in the milder FASD groups (groups 2-6) did not test positive for the episignature, which was expected given

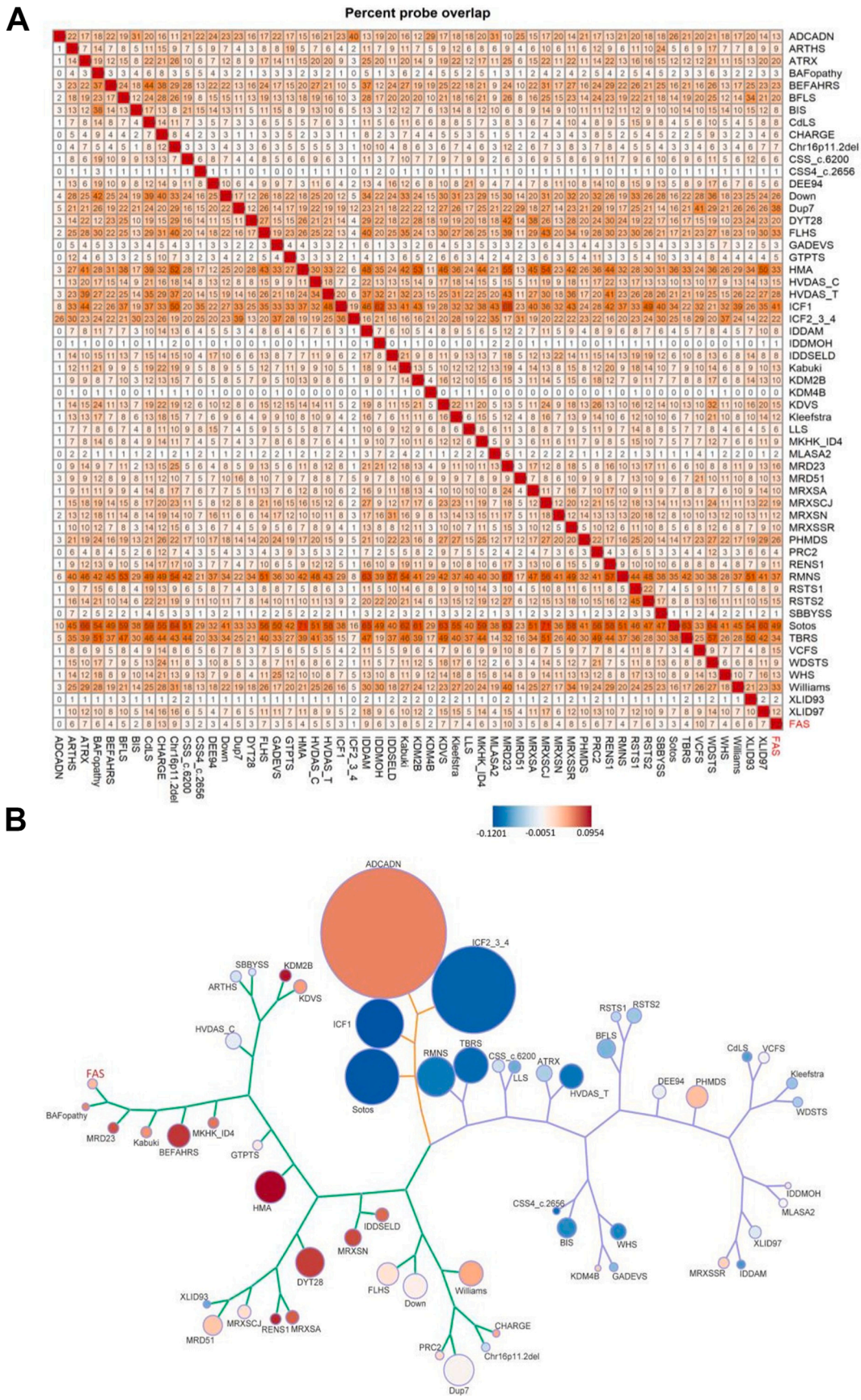

Figure 4 **Methylome profile comparison of FAS and EpiSign cohorts.** A. Heatmap showing the percentage of pairwise shared probes between cohorts. Below the diagonal, colors and values indicate the percentage of DMPs in cohorts along columns that are in common with

the phenotypic variability and the potential presence of additional molecular profiles in these subgroups. A positive classification indicates that the DNAm profile of an individual aligns with the FAS episignature, but it does not confirm a clinical diagnosis of FAS. Instead, it provides molecular evidence supporting a potential diagnosis when considered alongside clinical criteria. Our findings suggest that the episignature is highly robust for identifying individuals with confirmed FAS, but its sensitivity for milder FASD phenotypes is lower. This may indicate the presence of additional or distinct molecular profiles within the broader FASD spectrum. Importantly, we acknowledge that this approach aligns with standard practices in genomic diagnostics, in which biomarkers are used to support, rather than independently establish, a clinical diagnosis. Although genomic sequencing (ES/GS) is increasingly established as a tier-1 diagnostic test in many clinical settings, episignatures currently function as complementary tools that can provide additional diagnostic insights, particularly when sequencing results remain inconclusive (eg, VUS findings). The absence of a positive episignature for FASD in individuals with a high suspicion of FASD does not definitively exclude the possibility of prenatal alcohol exposure or the presence of FASD-related pathogenicity.[34] This may be due to multiple factors, including variability in the epigenetic response to alcohol exposure.[11,18,19] Additionally, it is possible that individuals with FASD who test negative for the episignature may fall into a distinct molecular subgroup that requires further investigation. Our observations of a general presence of hypermethylation in FAS cases compared with controls are in line with previous studies.[1,32-34] Moreover we detected a relatively high overlap of probes linked to the DNAm episignature of Sotos syndrome, a disorder known to affect growth and neurodevelopment. This suggests that some epigenetic disruptions in FASD may converge on pathways involved in neurodevelopmental regulation. Furthermore, both DMR analysis and expression quantitative trait methylation analyses corroborated some of our previous findings.[32] Although ES revealed a possible genetic etiology in only 1 case within our FAS cohort, this finding highlights the importance of integrating biomarkers to enhance diagnostic clarity. In this particular case, the individual carried a likely pathogenic variant in the *BRPF1* gene (HGNC: 14255, *BRPF1* c.3298C>T; p.Arg110*), a gene associated with intellectual disabilities, growth restriction, and craniofacial abnormalities, which may explain the elevated clinical diagnostic score. Notably, this same individual also tested positive for the FAS-specific DNAm episignature, suggesting a potential cooccurrence of FASD and a *BRPF1*-related disorder. Because no episignature has yet been reported for the condition associated with pathogenic variants in the *BRPF1* gene, we were unable, and it was beyond the scope of the present study, to train our FAS signature against a *BRPF1*-specific episignature. It is possible that such a signature will be identified in the future and, if appropriate, incorporated into future updates of our multiclass classifier. Moreover, this overlap underscores the complementary role of episignatures in genomic diagnostics. Although the episignature provides robust molecular support for a diagnosis of FAS, it is not intended to serve as a standalone tool. Instead, it should be interpreted alongside clinical findings and, where applicable, genetic data. As more individuals are screened using this combined approach, we anticipate further refinement of the episignature, potentially improving its sensitivity, especially in milder or genetically complex FASD cases.[47]

Sex differences in FASD-related outcomes have been widely reported, with males often exhibiting greater neurodevelopmental impairments, including hippocampal synaptic plasticity deficits[48,49] and higher rates of externalizing disorders.[50] In contrast, females more frequently present with internalizing symptoms and may demonstrate some neuroprotective mechanisms, such as increased glutamine synthetase expression.[48] It has been suggested that estrogen, in particular, may play a protective role in this context. In line with these observations, our cohort included more males than females. However, we did not observe any significant enrichment (FAS and FAS? t test, $P > .47$) of particular subcharacteristics of the 4-digit clinical diagnostic tool between males and females. Moreover, no significant sex differences were observed between cases testing positive or negative for the episignature. Because this study aimed to identify a general episignature for FASD, no further stratification was applied.

Previously published studies by our group and others also revealed a genome-wide DNAm change in individuals with FASD relative to healthy controls.[1,33,34,51] The later epigenome-wide association studies identified differential methylation patterns associated with FASD but did not demonstrate the ability to validate a biomarker that can be used to screen, as well as to differentiate, this relative to other episignature disorders. For example, Portales-Casamar et al[34] utilized the Illumina Human Methylation 450 array and examined 110 children with FASD, discovering 658 significantly differentially methylated sites that overlapped with 95 different genes between the children with FASD and the controls. A subsequent study by the same group validated many the previously identified DNAm changes.[34] Numerous other studies have similarly reported differential DNAm patterns in FASD,[51,52] highlighting the growing body of evidence in this field.

the cohorts along the rows. B. Tree and leaf plot visualization of Euclidean clustering of episignature groups using the top 500 DMPs of each cohort. Signature samples were aggregated and median value for each probe was calculated within each group. Cohorts are represented by nodes, in which the node color correlates to the overall mean difference in methylation of the cohort's DMPs, and the node size represents the relative size of the DMP set. Three main clusters were identified and are represented here by the colors of the edges.

However, our study, for the first time demonstrates the ability to use these methylation differences to develop an episignature biomarker.

The diagnosis of FASD relies on presence of 4 clinical features. Our study could not identify a direct association between the MVP and the 4-digit diagnostic code, whether considered individually or together (Supplemental Figure 2). To further refine FASD diagnostics and expand potential clinical application of the DNAm episignature, several research avenues could be pursued. For instance, exposure studies on model organisms can simulate prenatal alcohol exposure at different stages and doses, helping to understand its impact on DNAm and subsequent developmental outcomes.[53,54] Such studies can reveal critical windows of vulnerability and inform preventive strategies. Collection of patient cohorts with diverse phenotypes and different type of (gestational age) exposure is therefore crucial.

Efforts to develop screening tools for diagnosis of FASD have primarily focused on neurodevelopmental profiles or physiological markers. Although many classifiers have shown promise, challenges persist in creating models that are both accurate, easily reproducible, standardized and cost-effective, all of which can be achieved by use of molecular biomarker. Given the diverse presentation of FASD and its complex diagnosis, a critical direction for machine learning in this area involves integrating accessible modalities, such as neurodevelopmental assessment and facial imaging to improve sensitivity and specificity. Despite these advancements, FASD remains underdiagnosed, denying many children interventions such as specialized educational programs, behavioral therapies, and medical management that could significantly improve their quality of life.[2]

When examining the degree of overlap between FASD DMPs and other neurodevelopmental disorders, the highest observed overlap is 49% with Sotos syndrome. Sotos syndrome, caused by pathogenic variants in the *NSD1* gene (HGNC:14234), manifests with overgrowth, facial abnormalities, brain anomalies, seizures, and intellectual impairments.[55] Although Sotos syndrome has a large number of differentially methylated loci, and FASD is certainly not the only condition that shares a significant overlap with the Sotos syndrome signature, this overlap raises important questions regarding shared epigenetic pathways. Although the overlap of clinical features between Sotos syndrome and FASD is limited, for example, overgrowth and macrocephaly in Sotos compared with microcephaly and short stature in FASD, both conditions involve neurodevelopmental delay, highlighting the complexity of neurodevelopmental disorders. Such opposing phenotypes are not uncommon in epigenetically driven conditions. For instance, imprinting disorders such as Prader-Willi and Angelman syndromes are associated with opposing molecular aberrations yet share features such as learning difficulties, while displaying opposite phenotypes in areas such as food intake, muscle tone, and demeanor. Second, majority of changes observed in Sotos involve hypomethylation, whereas FAS exhibits predominantly hypermethylation, suggesting opposite-direction effects of epigenetic dysregulation, which might contribute to the contrasting growth patterns observed in these disorders. Although the 49% overlap of DMPs in the FAS signature with the Sotos signature may suggest shared systemic methylation abnormalities, such a claim requires further exploration because the opposite comparison does not show this effect, obviously due to the large number of DMPs in the Sotos signature. The latter effect is however also observed comparing DMP overlap in other epigenetic signatures.[22] Therefore, further in depth, cross-disorder studies are essential to understand how shared epigenetic modifications contribute to each disorder's phenotype and whether such methylation changes underlie broader neurodevelopmental pathways.

In this study, we identified in total 10 genomic regions expressing differentially methylated DNA that were also previously reported to be associated with FASD.[1,32] The previous reports partially included the same patients with FASD as this study. On the other hand, these studies differed in DNAm detection platform, as well as the DMR detection algorithm and statistical model designs, which are known factors that affect study outcomes. Therefore, the overlap of detected DMRs in those studies and this study, still supports the reliability of these 10 consistently associated DMRs. The overlapping DMRs, detected in this study and the aforementioned reports, were annotated to the following genes: *GLI2* (HGNC:4318), *TNFRSF19* (HGNC:11915), *DTNA* (HGNC:3057), *NECAB3* (HGNC:15851), *SEC61G* (HGNC:18277), REEP3 (HGNC:23711), *ZNF577* (HGNC:28673), *HNRNPF* (HGNC:5039), *MSC* (HGNC:7321), and *SDHAF1* (HGNC:33867). Notably, 2 of these DMRs, annotated to the *REEP3* and *SDHAF1* genes, were also reported to be functionally associated with in cis gene expression in patients with FASD but not in controls. REEP3 encoding receptor accessory protein 3, is involved in intracellular transport and mitotic spindle stability, key for normal cell division and neurodevelopment. Disruptions in mitotic spindle dynamics may impair neural progenitor cell division, plausibly contributing to FASD-associated brain abnormalities. The *SDHAF1* gene encodes for a LYR-motif protein and, to the best of our knowledge, has not previously been mechanistically linked to FASD in any report. Further investigation into its potential role in FASD pathogenesis may provide valuable insights into mitochondrial dysfunction and metabolic disturbances implicated in this disorder.[56,57]

One plausible causal mechanism for FASD is the that (prenatal) alcohol exposure interferes with the folate cycle and the 1-carbon metabolism, altering the availability of S-adenosylmethionine, the primary methyl donor for DNAm.[58,59] This can result in locus-specific epigenetic changes, especially in genes regulating early development. In parallel, ethanol-induced oxidative stress has been shown to promote both random and targeted DNAm alterations, including in neural crest cell lineages.[53,60,61] These cells give rise to craniofacial and neurodevelopment structures, which by definition are affected in FASD. Moreover, impaired

differentiation of neural crest cells, possibly driven by altered methylation of genes, such as the aforementioned *REEP3* gene, has been linked to a range of FASD phenotypes and overlaps mechanistically with Mendelian disorders involving mutations in epigenetic regulators.[32,58-60]

Because FAS cannot be confirmed through molecular analysis beyond the current DNAm signature, we recommend routinely requesting a detailed history of prenatal alcohol exposure, preferably including a complete 4-digit diagnostic code if available. In this context, DNAm episignature analysis may be considered as part of a broader diagnostic process for evaluating suspected of having FAS, alongside clinical assessments and complementary tests, such as neuropsychological and neuroimaging studies. Although a positive result represents a highly strong supporting evidence for molecular pathophysiology consistent with FAS, a negative episignature does not rule out the condition. This may be due to the presence of pathogenic variants in some individuals with FAS that contribute to the clinical presentation. Moreover, it cannot be excluded that such pathogenic variation may independently alter, or act in addition to, the FAS- DNAm profile.

Limitations

Several limitations should be considered when interpreting our findings. First, although the episignature demonstrates strong performance in confirmed FAS cases, representing the most extreme FASD phenotypes, it shows reduced sensitivity in individuals with milder or suspected FASD presentations, potentially leading to false negatives. Although our test cohort (revised version, groups 2-6) included patients who met the FAS 4-Digit Diagnostic Code criteria, these diagnoses were flagged because of, among other reasons, less reliable exposure histories or potential additional molecular profiles. Second, the FAS episignature is not intended as a standalone diagnostic tool; rather, it should be integrated with clinical evaluation and genetic testing to support diagnosis. Third, phenotypic and molecular heterogeneity within the FASD spectrum may also limit the episignature's sensitivity across all affected individuals. Some individuals who test negative may belong to distinct molecular subgroups not captured by the current signature, highlighting the need for further research. This further research may address timing, dose, and duration of prenatal alcohol exposure, as well as maternal factors such as age and nutrition.[2,62] In addition, genetic factors have been associated with alcohol metabolism and may mediate differences in teratogenic effects.[8,63-67] Fourth, the model was developed without stratification by sex, despite known sex-based differences in FASD outcomes, which may influence performance. In addition, the limited sample size and demographic diversity of the cohort may restrict generalizability, underscoring the need for larger and more diverse data sets. Moreover, because this study is based on cross-sectional data, the episignature's potential utility for monitoring disease progression or treatment response over time remains unexplored. Finally, reliance on peripheral blood samples may not fully capture brain-specific epigenetic changes relevant to the neurodevelopmental outcomes of FASD.

Conclusion

We demonstrated a relatively sensitive and specific DNAm episignature for patients with FAS. This episignature can be used as part of the diagnostic evidence in patients with clinically suspected FASD, or as part of a broader episignature classifier to rule out other clinically overlapping rare disorders with known episignature biomarkers. However, given the variability in sensitivity for milder phenotypes, future studies should explore multimodal approaches integrating neurodevelopmental and imaging data to improve diagnostic pathways for FAS.

Data Availability

Amsterdam UMC Raw gene expression sequencing data (fastq) are available at the European Genome-Phenome Archive: https://ega-archive.org/studies, under accession identifier EGAS00001006899, and Amsterdam UMC raw DNA methylation profiles are available at NCBI GeneExpression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo), under accession identifier GSE112987. Other data sets used in this study that are available publicly are previously described.[26] Anonymized data for each participant are described in the study. The individual genomic and epigenomic or any other personally identifiable data for other samples in the EpiSign Knowledge Database (EKD) are prohibited from deposition in publicly accessible databases because of institutional and ethics restrictions. Specifically, these include data and samples submitted from external institutions to London Health Sciences EKD that are participant to Institutional Material and Data Transfer agreements, data submitted to London Health Sciences for episignature assessment under Research Services Agreements, and research study cohorts under Institutional Research Ethics Approval (Western University REB 106302; and REB 116108). Some of the software packages used in this study are publicly available as described.

Acknowledgments

The cooperation of all the parents/care takers of the children with FASD included in this study is greatly acknowledged, as is the cooperation of the parents/care takers of the children in the healthy control group. The authors also thank Izabela M. Krzyzewska and Saskia M. Maas (Department of Human Genetics, Amsterdam UMC, Amsterdam, The Netherlands), Robert Smigiel (Department of Pediatrics, Endocrinology, Diabetology and Metabolic Diseases,

Wroclaw Medical University), and Jacek Niklinski and Monika A. Chomczyk (Department of Molecular Biology, Medical University of Bialystok, Białystok, Poland) for their help with the sample collection.

Funding

Funding for this study is provided in part by the Government of Canada through Genome Canada and the Ontario Genomics Institute (OGI-188).

Author Contributions

Conceptualization: M.A., M.M.A.M.M., M.M.v.H., I.V., E.F.T., B.S., P.H.; Data Curation: L.v.d.L., R.R., I.V., A.N.M., M.A.L., J.K., J.R.; Formal Analysis: L.v.d.L., R.R., M.A.L., J.K., J.R.; Investigation: M.A., A.L.-A., C.C.-A., A.M.C.-G., E.F.T., M.v.H., I.V.; Methodology: L.v.d.L., R.R., B.S., P.H.; Project Administration: L.v.d.L., J.K., B.S., P.H.; Supervision: B.S., P.H.; Validation: L.v.d.L., R.R., B.S., P.H.; Visualization: L.v.d.L., R.R., B.S., P.H.; Writing-original draft: L.v.d.L., R.R., B.S., P.H.; Writing-review and editing: L.v.d.L., B.S., P.H., I.V., E.F.T.

Ethics Declaration

Written informed consent was obtained from participants or participants family members before inclusion in this study. Consent included use of DNA and clinical information. In addition, this study was approved by the Western University Research Ethics Board (REB 106302, 10 August 2020) and The Medical Ethical Committee of the Amsterdam UMC, location AMC. EpiSign is a commercial software and is not publicly available.

ORCIDs

Peter Henneman: https://orcid.org/0000-0003-2179-7808

Conflict of Interest

Bekim Sadikovic is a shareholder in EpiSign Inc, a biotech firm involved in commercial application of EpiSign technology. All other authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2025.101586) contains supplemental material, which is available to authorized users.

Affiliations

[1]Department of Human Genetics, Amsterdam UMC, Amsterdam, The Netherlands; [2]Amsterdam Reproduction and Development, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands; [3]Verspeeten Clinical Genome Centre, London Health Science Centre, London, Canada; [4]Department of Clinical and Molecular Genetics, University Hospital Vall d'Hebron, Barcelona, Spain and Medicine Genetics Research Group, VHIR, Barcelona, Spain; [5]Emma Center for Personalized Medicine, Amsterdam UMC, Amsterdam, The Netherlands; [6]Department of Pediatrics, Amsterdam UMC, Amsterdam, The Netherlands; [7]Department of Pathology and Laboratory Medicine, Western University, London, Canada

References

1. Cobben JM, Krzyzewska IM, Venema A, et al. DNA methylation abundantly associates with fetal alcohol spectrum disorder and its subphenotypes. *Epigenomics*. 2019;11(7):767-785. http://doi.org/10.2217/epi-2018-0221
2. May PA, Baete A, Russo J, et al. Prevalence and characteristics of fetal alcohol spectrum disorders. *Pediatrics*. 2014;134(5):855-866. http://doi.org/10.1542/peds.2013-3319
3. Zarrei M, Hicks GG, Reynolds JN, et al. Copy number variation in fetal alcohol spectrum disorder. *Biochem Cell Biol*. 2018;96(2):161-166. http://doi.org/10.1139/bcb-2017-0241
4. Szakszon K, Lourenco CM, Callewaert BL, et al. Further delineation of the rare GDACCF (global developmental delay, absent or hypoplastic corpus callosum, dysmorphic facies syndrome): genotype and phenotype of 22 patients with *ZNF148* mutations. *J Med Genet*. 2024;61(2):132-141. http://doi.org/10.1136/jmg-2022-109030
5. Tan CA, Topper S, Del Gaudio D, et al. Characterization of patients referred for non-specific intellectual disability testing: the importance of autosomal genes for diagnosis. *Clin Genet*. 2016;89(4):478-483. http://doi.org/10.1111/cge.12575
6. Astley SJ, Clarren SK. Diagnosing the full spectrum of fetal alcohol-exposed individuals: introducing the 4-digit diagnostic code. *Alcohol Alcohol*. 2000;35(4):400-410. http://doi.org/10.1093/alcalc/35.4.400
7. Wozniak JR, Riley EP, Charness ME. Clinical presentation, diagnosis, and management of fetal alcohol spectrum disorder. *Lancet Neurol*. 2019;18(8):760-770. http://doi.org/10.1016/S1474-4422(19)30150-4
8. Sambo D, Goldman D. Genetic influences on fetal alcohol spectrum disorder. *Genes (Basel)*. 2023;14(1):195. http://doi.org/10.3390/genes14010195
9. Carter RC, Yang Z, Akkaya-Hocagil T, et al. Genetic admixture predictors of fetal alcohol spectrum disorders (FASD) in a South African population. *Gene*. 2024;931:148854. http://doi.org/10.1016/j.gene.2024.148854
10. Eberhart JK, Parnell SE. The genetics of fetal alcohol spectrum disorders. *Alcohol Clin Exp Res*. 2016;40(6):1154-1165. http://doi.org/10.1111/acer.13066
11. May PA, Blankenship J, Marais AS, et al. Approaching the prevalence of the full spectrum of fetal alcohol spectrum disorders in a South African population-based study. *Alcohol Clin Exp Res*. 2013;37(5):818-830. http://doi.org/10.1111/acer.12033
12. Benz J, Rasmussen C, Andrew G. Diagnosing fetal alcohol spectrum disorder: history, challenges and future directions. *Paediatr Child Health*. 2009;14(4):231-237. http://doi.org/10.1093/pch/14.4.231

13. Jonsson E. Fetal alcohol spectrum disorders (FASD): a policy perspective. *Can J Psychiatry*. 2019;64(3):161-163. http://doi.org/10.1177/0706743718773706
14. Običan S, Scialli AR. Teratogenic exposures. *Am J Med Genet C Semin Med Genet*. 2011;157C(3):150-169. http://doi.org/10.1002/ajmg.c.30310
15. Roozen S, Peters GJ, Kok G, Townend D, Nijhuis J, Curfs L. Worldwide prevalence of fetal alcohol spectrum disorders: a systematic literature review including meta-analysis. *Alcohol Clin Exp Res*. 2016;40(1):18-32. http://doi.org/10.1111/acer.12939
16. Popova S, Charness ME, Burd L, et al. Fetal alcohol spectrum disorders. *Nat Rev Dis Primers*. 2023;9(1):11. http://doi.org/10.1038/s41572-023-00420-x
17. Murawski NJ, Moore EM, Thomas JD, Riley EP. Advances in diagnosis and treatment of fetal alcohol spectrum disorders: from animal models to human studies. *Alcohol Res*. 2015;37(1):97-108.
18. Breuer L, Greenmyer JR, Wilson T. Clinical diagnosis and management of fetal alcohol spectrum disorder and sensory processing disorder in children. *Children (Basel)*. 2024;11(1):108. http://doi.org/10.3390/children11010108
19. Petrenko CL, Alto ME. Interventions in fetal alcohol spectrum disorders: an international perspective. *Eur J Med Genet*. 2017;60(1):79-91. http://doi.org/10.1016/j.ejmg.2016.10.005
20. Dozet D, Burd L, Popova S. Screening for alcohol use in pregnancy: a review of current practices and perspectives. *Int J Ment Health Addict*. 2023;21(2):1220-1239. http://doi.org/10.1007/s11469-021-00655-3
21. Chasnoff IJ, Wells AM, King L. Misdiagnosis and missed diagnoses in foster and adopted children with prenatal alcohol exposure. *Pediatrics*. 2015;135(2):264-270. http://doi.org/10.1542/peds.2014-2171
22. Levy MA, McConkey H, Kerkhof J, et al. Novel diagnostic DNA methylation episignatures expand and refine the epigenetic landscapes of Mendelian disorders. *HGG Adv*. 2021;3(1):100075. http://doi.org/10.1016/j.xhgg.2021.100075
23. Acharya A, Kavus H, Dunn P, et al. Delineating the genotypic and phenotypic spectrum of *HECW2*-related neurodevelopmental disorders. *J Med Genet*. 2022;59(7):669-677. http://doi.org/10.1136/jmedgenet-2021-107871
24. Aref-Eshghi E, Bend EG, Colaiacovo S, et al. Diagnostic utility of genome-wide DNA methylation testing in genetically unsolved individuals with suspected hereditary conditions. *Am J Hum Genet*. 2019;104(4):685-700. http://doi.org/10.1016/j.ajhg.2019.03.008
25. Aref-Eshghi E, Kerkhof J, Pedro VP, et al. Evaluation of DNA methylation episignatures for diagnosis and phenotype correlations in 42 Mendelian neurodevelopmental disorders. *Am J Hum Genet*. 2021;108(6):1161-1163. http://doi.org/10.1016/j.ajhg.2021.04.022
26. Aref-Eshghi E, Rodenhiser DI, Schenkel LC, et al. Genomic DNA methylation signatures enable concurrent diagnosis and clinical genetic variant classification in neurodevelopmental syndromes. *Am J Hum Genet*. 2018;102(1):156-174. http://doi.org/10.1016/j.ajhg.2017.12.008
27. Kerkhof J, Squeo GM, McConkey H, et al. DNA methylation episignature testing improves molecular diagnosis of Mendelian chromatinopathies. *Genet Med*. 2022;24(1):51-60. http://doi.org/10.1016/j.gim.2021.08.007
28. Sadikovic B, Levy MA, Kerkhof J, et al. Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders. *Genet Med*. 2021;23(6):1065-1074. http://doi.org/10.1038/s41436-020-01096-4
29. Kerkhof J, Rastin C, Levy MA, et al. Diagnostic utility and reporting recommendations for clinical DNA methylation episignature testing in genetically undiagnosed rare diseases. *Genet Med*. 2024;26(5):101075. http://doi.org/10.1016/j.gim.2024.101075
30. Haghshenas S, Karimi K, Stevenson RE, et al. Identification of a DNA methylation episignature for recurrent constellations of embryonic malformations. *Am J Hum Genet*. 2024;111(8):1643-1655. http://doi.org/10.1016/j.ajhg.2024.07.005
31. Haghshenas S, Putoux A, Reilly J, et al. Discovery of DNA methylation signature in the peripheral blood of individuals with history of antenatal exposure to valproic acid. *Genet Med*. 2024;26(10):101226. http://doi.org/10.1016/j.gim.2024.101226
32. Krzyzewska IM, Lauffer P, Mul AN, et al. Expression quantitative trait methylation analysis identifies whole blood molecular footprint in Fetal Alcohol Spectrum Disorder (FASD). *Int J Mol Sci*. 2023;24(7):6601. http://doi.org/10.3390/ijms24076601
33. Lussier AA, Morin AM, MacIsaac JL, et al. DNA methylation as a predictor of fetal alcohol spectrum disorder. *Clin Epigenet*. 2018;10:5. http://doi.org/10.1186/s13148-018-0439-6
34. Portales-Casamar E, Lussier AA, Jones MJ, et al. DNA methylation signature of human fetal alcohol spectrum disorder. *Epigenetics Chromatin*. 2016;9:25. http://doi.org/10.1186/s13072-016-0074-4
35. Zhou W, Triche TJ Jr, Laird PW, Shen H. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res*. 2018;46(20):e123. http://doi.org/10.1093/nar/gky691
36. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res*. 2017;45(4):e22. http://doi.org/10.1093/nar/gkw967
37. Aref-Eshghi E, Kerkhof J, Pedro VP, et al. Evaluation of DNA methylation episignatures for diagnosis and phenotype correlations in 42 Mendelian neurodevelopmental disorders. *Am J Hum Genet*. 2020;106(3):356-370. http://doi.org/10.1016/j.ajhg.2020.01.019
38. Chen YA, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8(2):203-209. http://doi.org/10.4161/epi.23470
39. Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*. 2016;17(1):208. http://doi.org/10.1186/s13059-016-1066-1
40. Ho DE, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference. *J Stat Softw*. 2011;42(8):1-28. http://doi.org/10.18637/jss.v042.i08
41. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. http://doi.org/10.1093/nar/gkv007
42. Levy MA, Relator R, McConkey H, et al. Functional correlation of genome-wide DNA methylation profiles in genetic neurodevelopmental disorders. *Hum Mutat*. 2022;43(11):1609-1628. http://doi.org/10.1002/humu.24446
43. Peters TJ, Buckley MJ, Statham AL, et al. De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin*. 2015;8:6. http://doi.org/10.1186/1756-8935-8-6
44. Cavalcante RG, Sartor MA. annotatr: genomic regions in context. *Bioinformatics*. 2017;33(15):2381-2383. http://doi.org/10.1093/bioinformatics/btx183
45. Girirajan S, Elsas LJ 2nd, Devriendt K, Elsea SH. RAI1 variations in Smith-Magenis syndrome patients without 17p11.2 deletions. *J Med Genet*. 2005;42(11):820-828. http://doi.org/10.1136/jmg.2005.031211
46. Dharmadhikari AV, Kang SH, Szafranski P, et al. Small rare recurrent deletions and reciprocal duplications in 2q21.1, including brain-specific ARHGEF4 and GPR148. *Hum Mol Genet*. 2012;21(15):3345-3355. http://doi.org/10.1093/hmg/dds166
47. Yan K, Rousseau J, Littlejohn RO, et al. Mutations in the chromatin regulator gene BRPF1 cause syndromic intellectual disability and deficient histone acetylation. *Am J Hum Genet*. 2017;100(1):91-104. http://doi.org/10.1016/j.ajhg.2016.11.011
48. Sickmann HM, Patten AR, Morch K, et al. Prenatal ethanol exposure has sex-specific effects on hippocampal long-term potentiation. *Hippocampus*. 2014;24(1):54-64. http://doi.org/10.1002/hipo.22203
49. Fontaine CJ, Patten AR, Sickmann HM, Helfer JL, Christie BR. Effects of pre-natal alcohol exposure on hippocampal synaptic plasticity: sex, age and methodological considerations. *Neurosci Biobehav Rev*. 2016;64:12-34. http://doi.org/10.1016/j.neubiorev.2016.02.014
50. Flannigan K, Poole N, Cook J, Unsworth K. Sex-related differences among individuals assessed for fetal alcohol spectrum disorder in

Canada. *Alcohol Clin Exp Res (Hoboken)*. 2023;47(3):613-623. http://doi.org/10.1111/acer.15017
51. Suttie M, Kable J, Mahnke AH, Bandoli G. Machine learning approaches to the identification of children affected by prenatal alcohol exposure: a narrative review. *Alcohol Clin Exp Res (Hoboken)*. 2024;48(4):585-595. http://doi.org/10.1111/acer.15271
52. Okazaki S, Otsuka I, Shinko Y, et al. Epigenetic clock analysis in children with fetal alcohol spectrum disorder. *Alcohol Clin Exp Res*. 2021;45(2):329-337. http://doi.org/10.1111/acer.14532
53. Leung ECH, Jain P, Michealson MA, Choi H, Ellsworth-Kopkowski A, Valenzuela CF. Recent breakthroughs in understanding the cerebellum's role in fetal alcohol spectrum disorder: a systematic review. *Alcohol*. 2024;119:37-71. http://doi.org/10.1016/j.alcohol.2023.12.003
54. Clark CA, Nakhid D, Baldwin-Oneill G, et al. Prevalence of co-occurring diagnoses in people exposed to alcohol prenatally: findings from a meta-analysis. *J Affect Disord*. 2024;358:163-174. http://doi.org/10.1016/j.jad.2024.05.035
55. Choufani S, Cytrynbaum C, Chung BH, et al. NSD1 mutations generate a genome-wide DNA methylation signature. *Nat Commun*. 2015;6:10207. http://doi.org/10.1038/ncomms10207
56. Jain-Ghai S, Cameron JM, Al Maawali A, et al. Complex II deficiency–a case report and review of the literature. *Am J Med Genet A*. 2013;161A(2):285-294. http://doi.org/10.1002/ajmg.a.35714
57. Ghezzi D, Goffrini P, Uziel G, et al. SDHAF1, encoding a LYR complex-II specific assembly factor, is mutated in SDH-defective infantile leukoencephalopathy. *Nat Genet*. 2009;41(6):654-656. http://doi.org/10.1038/ng.378
58. McKay JA, Williams EA, Mathers JC. Folate and DNA methylation during in utero development and aging. *Biochem Soc Trans*. 2004;32 (Pt 6):1006-1007. http://doi.org/10.1042/BST0321006
59. Nickels EM, Li S, Morimoto L, et al. Periconceptional folate intake influences DNA methylation at birth based on dietary source in an analysis of pediatric acute lymphoblastic leukemia cases and controls. *Am J Clin Nutr*. 2022;116(6):1553-1564. http://doi.org/10.1093/ajcn/nqac283
60. Chen SY, Kannan M. Neural crest cells and fetal alcohol spectrum disorders: mechanisms and potential targets for prevention. *Pharmacol Res*. 2023;194:106855. http://doi.org/10.1016/j.phrs.2023.106855
61. Kietzman HW, Everson JL, Sulik KK, Lipinski RJ. The teratogenic effects of prenatal ethanol exposure are exacerbated by Sonic Hedgehog or GLI2 haploinsufficiency in the mouse. *PLoS One*. 2014;9(2):e89448. http://doi.org/10.1371/journal.pone.0089448
62. Young JK, Giesbrecht HE, Eskin MN, Aliani M, Suh M. Nutrition implications for fetal alcohol spectrum disorder. *Adv Nutr*. 2014;5 (6):675-692. http://doi.org/10.3945/an.113.004846
63. Chasnoff IJ. Fetal alcohol syndrome in twin pregnancy. *Acta Genet Med Gemellol (Roma)*. 1985;34(3-4):229-232. http://doi.org/10.1017/s0001566000004797
64. Christoffel KK, Salafsky I. Fetal alcohol syndrome in dizygotic twins. *J Pediatr*. 1975;87(6 Pt 1):963-967. http://doi.org/10.1016/s0022-3476(75)80919-x
65. Riese ML. Maternal alcohol and pentazocine abuse: neonatal behavior and morphology in an opposite-sex twin pair. *Acta Genet Med Gemellol (Roma)*. 1989;38(1-2):49-56. http://doi.org/10.1017/s0001566000002828
66. Riikonen RS. Difference in susceptibility to teratogenic effects of alcohol in discordant twins exposed to alcohol during the second half of gestation. *Pediatr Neurol*. 1994;11(4):332-336. http://doi.org/10.1016/0887-8994(94)90012-4
67. Streissguth AP, Dehaene P. Fetal alcohol syndrome in twins of alcoholic mothers: concordance of diagnosis and IQ. *Am J Med Genet*. 1993;47(6):857-861. http://doi.org/10.1002/ajmg.1320470612