

Washington University School of Medicine

Digital Commons@Becker

---

2020-Current year OA Pubs

Open Access Publications

---

6-1-2021

## Clinical epigenomics: Genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders

Bekim Sadikovic

Jorge L Granadillo

et al.

Follow this and additional works at: [https://digitalcommons.wustl.edu/oa\\_4](https://digitalcommons.wustl.edu/oa_4)






Part of the [Medicine and Health Sciences Commons](#)

---



## ARTICLE

## Clinical epigenomics: genome-wide DNA methylation analysis for the diagnosis of Mendelian disorders

Bekim Sadikovic<sup>1,2</sup> , Michael A. Levy<sup>1,2</sup>, Jennifer Kerkhof<sup>1,2</sup>, Erfan Aref-Eshghi<sup>1,2</sup>, Laila Schenkel<sup>1,2</sup>, Alan Stuart<sup>1,2</sup>, Haley McConkey<sup>1,2</sup>, Peter Henneman<sup>3</sup>, Andrea Venema<sup>3</sup>, Charles E. Schwartz<sup>4</sup>, Roger E. Stevenson<sup>4</sup>, Steven A. Skinner<sup>4</sup>, Barbara R. DuPont<sup>4</sup>, Robin S. Fletcher<sup>4</sup>, Tugce B. Balci<sup>5,6</sup>, Victoria Mok Siu<sup>5,6</sup>, Jorge L. Granadillo<sup>7</sup>, Jennefer Masters<sup>1,2</sup>, Mike Kadour<sup>1,2</sup>, Michael J. Friez<sup>4</sup>, Mieke M. van Haelst<sup>3</sup>, Marcel M. A. M. Mannens<sup>3</sup>, Raymond J. Louie<sup>4</sup>, Jennifer A. Lee<sup>4</sup>, Matthew L. Tedder<sup>4</sup>  and Marielle Alders<sup>3</sup> 

**PURPOSE:** We describe the clinical implementation of genome-wide DNA methylation analysis in rare disorders across the EpiSign diagnostic laboratory network and the assessment of results and clinical impact in the first subjects tested.

**METHODS:** We outline the logistics and data flow between an integrated network of clinical diagnostics laboratories in Europe, the United States, and Canada. We describe the clinical validation of EpiSign using 211 specimens and assess the test performance and diagnostic yield in the first 207 subjects tested involving two patient subgroups: the targeted cohort (subjects with previous ambiguous/inconclusive genetic findings including genetic variants of unknown clinical significance) and the screening cohort (subjects with clinical findings consistent with hereditary neurodevelopmental syndromes and no previous conclusive genetic findings).

**RESULTS:** Among the 207 subjects tested, 57 (27.6%) were positive for a diagnostic episinature including 48/136 (35.3%) in the targeted cohort and 8/71 (11.3%) in the screening cohort, with 4/207 (1.9%) remaining inconclusive after EpiSign analysis.

**CONCLUSION:** This study describes the implementation of diagnostic clinical genomic DNA methylation testing in patients with rare disorders. It provides strong evidence of clinical utility of EpiSign analysis, including the ability to provide conclusive findings in the majority of subjects tested.

*Genetics in Medicine* (2021) 23:1065–1074; <https://doi.org/10.1038/s41436-020-01096-4>

## INTRODUCTION

Mendelian disorders are estimated to occur at a rate of 40 to 82 per 1,000 live births.<sup>1</sup> However, if all congenital anomalies are considered, approximately 8% of individuals are estimated to have a genetic disorder before reaching adulthood.<sup>2</sup> Clinical presentations in most genetic disorders include developmental delay and intellectual disability (DD/ID), sometimes in combination with other features including dysmorphism, neuromuscular phenotypes, and other systemic constellations of syndromes.<sup>3</sup> Despite rapid advances in our understanding of the human genome, nearly two-thirds of the patients with suspected rare genetic disorders remain without a conclusive molecular genetic diagnosis.<sup>4</sup>

Evolution of genetic testing from single-nucleotide assessment to clinical exome and genome sequencing, while increasing the diagnostic yield to an average of 36%,<sup>4</sup> has also resulted in a significant increase in ambiguous or uncertain genetic findings, referred to as variants of unknown clinical significance (VUS). Despite concerted efforts to standardize guidelines for the interpretation of sequence variants<sup>5</sup> and to define the functional evidence for variant classification,<sup>6</sup> a large proportion of VUS remain without conclusive clinical interpretation. Also, the understanding of the impact of genetic variation outside of protein-coding DNA sequences is very limited, and as such, the majority of genetic testing in clinical laboratories is focused on exonic and

short surrounding intronic sequences. Family variant cosegregation studies, in silico prediction algorithms, and gene-specific functional studies may help resolve VUS findings, but in the majority of cases these are not available, feasible, or conclusive.

One functional consequence of genetic defects in patients with hereditary neurodevelopmental disorders is the disruption of genomic DNA methylation.<sup>7</sup> DNA methylation is an epigenetic modification, resulting in changes in structural and chemical properties of the DNA, impacting molecular mechanisms including chromatin assembly and gene transcription.<sup>8</sup> Our group and others have demonstrated that individuals among a growing number of rare disorders exhibit DNA methylation “episignatures” or “EpiSigns” as highly sensitive and specific DNA methylation biomarkers.<sup>9–22</sup> These genome-wide DNA methylation profiles currently include over 40 rare disorders in association with more than 60 genes, and can be gene domain, gene level, as well as protein complex specific. DNA methylation episignatures are detectable in peripheral blood and are highly sensitive and specific for each disorder. As such, they represent effective biomarkers for the testing of patients with a broadening range of neurodevelopmental genetic conditions, as well as a reflex functional test for patients with ambiguous genetic test findings or clinical phenotypes.<sup>23</sup>

Genomic DNA methylation analysis is also adaptable to the routine analytical processes in clinical laboratories. Cytosine

<sup>1</sup>Molecular Genetics Laboratory, Molecular Diagnostics Division, London Health Sciences Centre, London, ON, Canada. <sup>2</sup>Department of Pathology and Laboratory Medicine, Western University, London, ON, Canada. <sup>3</sup>Amsterdam University Medical Center, University of Amsterdam, Department of Clinical Genetics, Amsterdam Reproduction and Development Research Institute, Amsterdam, The Netherlands. <sup>4</sup>Greenwood Genetic Center, Greenwood, SC, USA. <sup>5</sup>Department of Pediatrics, Division of Medical Genetics, Western University, London, ON, Canada. <sup>6</sup>Medical Genetics Program of Southwestern Ontario, London Health Sciences Centre, London, ON, Canada. <sup>7</sup>Division of Genetics and Genomic Medicine, Department of Pediatrics, Washington University School of Medicine in St. Louis, St. Louis, MO, USA. <sup>✉</sup>email: Bekim.Sadikovic@lhsc.on.ca; mtedder@ggc.org; m.alders@amsterdamumc.nl

methylation is a highly stable analyte, and genome-wide DNA methylation data can be generated on a microarray platform. In parallel with episinature screening, genomic DNA methylation analysis enables concurrent and highly sensitive and specific assessment of imprinting disorders<sup>24</sup> and fragile X syndrome,<sup>25</sup> enabling further test consolidation in this patient population. One key technical challenge in the clinical setting is data analysis. This requires the development of large-scale reference DNA methylation databases, including disorder and tissue-specific reference data sets and controls, and sophisticated analytical processes including machine learning algorithms as analytical classifiers. Testing must be performed in a regulated clinically certified environment, with adherence to the required quality management procedures, and clinical quality metrics, all under professional clinical oversight.

In this study we describe the implementation and validation of EpiSign, a clinical genome-wide DNA methylation test for patients with rare Mendelian disorders, based on the Illumina Infinium methylation array technology and the EpiSign Knowledge Database (EKD). We describe the quality metrics and clinical validation metrics within an integrated network of licensed academic nonprofit clinical laboratories in Europe, Canada, and the United States. We describe the clinical performance and the diagnostic yield in subjects tested between initiation of the service in November 2019 to June 2020. This study demonstrates the clinical utility of genomic DNA methylation testing in patients with Mendelian neurodevelopmental disorders.

## MATERIALS AND METHODS

### Clinical validation and patient cohorts

The validation cohort (Table S1) was designed to clinically validate and assess quality metrics of the EpiSign test across the EpiSign diagnostic laboratory network, and consisted of 211 archived peripheral blood DNA samples including samples with confirmed diagnosis of one of 43 genetic syndromes included in the EpiSign v2 genome-wide DNA methylation assay (Table 1), or controls. The genetic variation in these specimens were classified as pathogenic or likely pathogenic based on the American College of Medical Genetics and Genomics (ACMG) guidelines for interpretation of genomic sequence variants.<sup>5</sup> Technical inter and intrarun replicates were assessed for concordance based on methylation variant pathogenicity (MVP) score (within 0.05) and clustering analysis. The clinical testing cohort (Table S2) consists of peripheral blood DNA samples from 207 subjects, referred by physicians based on individual clinical discretion, who have received clinical EpiSign testing. All subjects provided informed consent for clinical genetic testing as part of pretest counseling.

DNA methylation analysis was performed using the Illumina Infinium EPIC bead chip arrays as previously described<sup>9</sup> by the clinical testing laboratories: Greenwood Genetic Center (Greenwood, SC, USA) and Amsterdam University Medical Center and partner labs (Amsterdam, Netherlands). Data from validation and clinical testing specimens (November 2019 and June 2020) were blinded and submitted to the clinical bioinformatics laboratory (Molecular Diagnostics Laboratory, London Health Sciences Centre, Western University, London, Canada) through a secure file transfer protocol and housed on the hospital clinical encrypted servers.

### DNA methylation data analysis

Analysis of the DNA methylation array data was performed by the clinical bioinformatics laboratory using Illumina Infinium EPIC arrays. Methylation data for each sample were compared to the established DNA methylation episinatures for the 43 disorders (Table 1) which are part of the EpiSign clinical test. EpiSign analysis utilized the EKD, a clinical database with >5,000 peripheral blood DNA methylation profiles including disorder-specific reference cohorts and normal (general population samples with various age and racial backgrounds) controls housed at London Health Sciences Centre Molecular Diagnostics Laboratory (<https://www.lhsc.on.ca/palm/molecular.html>). Individual DNA methylation data for each subject were compared with the EKD using the support vector machine (SVM) based classification algorithm for EpiSign disorders. Methylation variant Pathogenicity (MVP) score is generated ranging between 0 and 1,

representing the confidence of prediction for the specific class the SVM was trained to detect. Conversion of SVM decision values to these scores was carried out according to the Platt scaling method.<sup>26</sup> Classification for a specific EpiSign disorder included MVP score assessment with a general threshold of >0.5 for positive, <0.1 negative, 0.1–0.5 inconclusive or low confidence, hierarchical clustering and the multidimensional scaling (MDS) of subject's methylation data relative to the disorder-specific EpiSign probe sets and controls. A detailed description of this analytics protocol was described previously.<sup>9,27</sup>

### Clinical assessment and reporting

DNA methylation analysis results were clinically verified by a board-certified clinical molecular geneticist at the clinical bioinformatics laboratory. Result categories include positive (matching an EpiSign disorder), negative (not matching any EpiSign disorder), and inconclusive (described in detail in results). The report is then reviewed and verified by a board-certified laboratory professional in the clinical testing laboratory, and a clinical report describing the EpiSign results (positive, negative, or inconclusive for a particular episinature) is issued to the requesting physician.

## RESULTS

### EpiSign clinical validation

The 211 validation specimens are described in Table S1. Data were generated at the clinical testing laboratories, anonymized, and submitted to the clinical bioinformatics laboratory for EpiSign analysis. EpiSign analysis was concordant with the previous genetic findings in 207/211 samples. A positive control cohort included 143 samples with various genetic disorders with previously reported DNA episinatures, imprinting and uniparental disomy disorders, and fragile X syndrome. Of these, 139 were concordant for the expected episinature. The discordant cases included a subject (Val118) with a previously reported likely pathogenic variant, *KANSL1* (NM\_001193466.1): c.297\_307del; p. Gly100Glnfs\*6, related to Koolen–de Vries syndrome (KDVS), and another subject (Val26) with a previously reported likely pathogenic variant, *CREBBP* (NM\_004380.2): c.4480C>A; p. Pro1494Thr, related to the Rubinstein–Taybi syndrome (RSTS). In the case Val118 with the *KANSL1* likely pathogenic variant, and clinical features consistent with KDVS, the MVP score for KDVS was zero. Some individuals carry a duplication harboring part of the *KANSL1* gene elsewhere in the genome, which may cause variants in the duplicated region to be erroneously assigned to the *KANSL1* gene. However, assessment of exome sequencing data, array comparative genomic hybridization (array CGH), and multiplex ligation-dependent probe amplification (MLPA) analysis showed no evidence of this this duplication. The MVP score in the second case, Val26, with *CREBBP*(NM\_004380.2):c.4480C>A, p. (Pro1494Thr), for RSTS was zero. Potential mosaicism of the *CREBBP* variant was not apparent. The remaining two discordant cases (Val128 and Val140) both had previously identified *SMARCC2* pathogenic variants. Fifty-six samples were normal reference controls (NC) and were all concordant. The remaining 12 samples (other controls in Table S1) included samples with previous genetic findings that are currently undetectable by EpiSign analysis including samples with low-level mosaicism for imprinting disorders, fragile X female heterozygotes, 16p11del and *ARID2* pathogenic variant. There were 10 interrun duplicates (samples processed in different array batches) and 18 intrarun duplicates (replicate samples processed in the same batch), and all were concordant. EpiSign v2 validation included 55 replicates of samples used in v1 validation and all were concordant.

### EpiSign clinical testing

EpiSign analysis includes genetic conditions that exhibit DNA methylation episinatures as well as the common imprinting disorders and fragile X (Table 1). A total of 207 subjects were

**Table 1.** Disorders detectable by EpiSign v2.

Disease/disorder	Causative gene(s)/region	New in EpiSign V2	Validation cohort positives	Clinical cohort positives
$\alpha$ -thalassemia/mental retardation X-linked syndrome (ATR-X)	<i>ATRX</i> (301404)	No	2	4
Autism, susceptibility to, 18 (AUT18)	<i>CHD8</i> (610528)	Yes	6	3
BAFopathies: Coffin–Siris (CSS1–4, and 8) and Nicolaides–Baraitser (NCBRS) syndromes	<i>ARID1B</i> (135900), <i>ARID1A</i> (614607), <i>SMARCB1</i> (614608), <i>SMARCA4</i> (614609), <i>SMARCC2</i> (618362), <i>SMARCA2</i> (601358)	<i>ARID1A</i> , <i>SMARCC2</i>	28	10
Blepharophimosis intellectual disability syndrome, <i>SMARCA2</i> type	<i>SMARCA2</i> (OMIM not available, PMID: 32694869)	Yes	2	
Börjeson–Forssman–Lehmann syndrome (BFLS)	<i>PHF6</i> (301900)	Yes	1	
Cerebellar ataxia, deafness, and narcolepsy, autosomal dominant (ADCADN)	<i>DNMT1</i> (604121)	No		
CHARGE syndrome	<i>CHD7</i> (214800)	No	6	4
Cornelia de Lange syndrome (CdLS)	<i>NIPBL</i> (122470), <i>RAD21</i> (614701), <i>SMC3</i> (610759), <i>SMC1A</i> (300590)	No	11	6
Down syndrome	Trisomy 21 (190685)	No	1	
Epileptic encephalopathy, childhood-onset (EEOC)	<i>CHD2</i> (615369)	Yes	1	
Floating Harbor syndrome (FLHS)	<i>SRCAP</i> (136140)	No		
Genitopatellar syndrome (GTPTS) and Ohdo syndrome, <i>SBBYSS</i> variant ( <i>SBBYSS</i> )	<i>KAT6B</i> (606170; 603736)	No	3	2
Helsmoortel–van der Aa syndrome (HVDAS) <sup>a</sup>	<i>ADNP</i> (615873)	No	9	2
Hunter–McAlpine syndrome (HMA)	Chr5q35-qter duplication (601379)	Yes		
Immunodeficiency–centromeric instability–facial anomalies syndrome (ICF) <sup>b</sup>	<i>DNMT3B</i> (242860), <i>CDCA7</i> (616910), <i>ZBTB24</i> (614069), <i>HELLS</i> (616911)	Yes		
Kabuki syndrome	<i>KMT2D</i> (147920), <i>KDM6A</i> (300867)	<i>KDM6A</i>	12	5
Kleefstra syndrome	<i>EHMT1</i> (610253)	Yes	2	2
Koolen–de Vries syndrome (KDVS)	<i>KANSL1</i> (610443)	Yes	5	
Mental retardation, autosomal dominant 23 (MRD23)	<i>SETD5</i> (615761)	Yes	4	
Mental retardation, autosomal dominant 51 (MRD51)	<i>KMT5B</i> (617788)	Yes		
Mental retardation, X-linked 93 (MRD93)	<i>BRWD3</i> (300659)	Yes	1	1
Mental retardation, X-linked 97 (MRD97)	<i>ZNF711</i> (300803)	Yes		1
Mental retardation, X-linked syndromic, Nascimento-type (MRXSN)	<i>UBE2A</i> (300860)	Yes	1	
Mental retardation, X-linked, Snyder–Robinson type (MRXSSR)	<i>SMS</i> (309583)	Yes		1
Mental retardation, X-linked, syndromic, Claes–Jensen syndrome (MRXSCJ)	<i>KDM5C</i> (300534)	No	6	1
PCR2 complex (Weaver (WVS) and Cohen–Gibson (COGIS))	<i>EZH2</i> (277590), <i>EED</i> (617561)	Yes		
Rahman syndrome (RMNS)	<i>HIST1H1E</i> (617537)	Yes		
Rubinstein–Taybi syndrome (RSTS)	<i>CREBBP</i> (180849), <i>EP300</i> (613684)	Yes	2	3
SETD1B-related syndrome	<i>SETD1B</i> (619000)	Yes		1
Sotos syndrome	<i>NSD1</i> (117500)	No	8	4
Tatton–Brown–Rahman syndrome (TBRS)	<i>DNMT3A</i> (615879)	Yes	1	1
Wiedemann–Steiner syndrome (WDSTS)	<i>KMT2A</i> (605130)	Yes	3	2
Williams–Beuren deletion syndrome (WBS) and Williams–Beuren regions duplication syndrome (Dup7) <sup>c</sup>	7q11.23 deletion (194050)/ duplication (609757)	No	2	
Wolf–Hirschhorn syndrome (WHS)	Chr4p16.13 deletion (194190)	Yes		
Fragile X syndrome (FXS)	<i>TNR/FMR1</i> (300624)	No	6	
Mental retardation, FRA12A type	<i>TNR/DIP2B</i> (136630)	No		1

**Table 1** continued

Disease/disorder	Causative gene(s)/region	New in EpiSign V2	Validation cohort positives	Clinical cohort positives
Angelman syndrome (AS)	ID/ <i>UBE3A</i> (105830)	No	6	1
Prader–Willi syndrome	ID/15q11 ( <i>SNRPN</i> , <i>NDN</i> ) (176270)	No	3	
Silver–Russell syndrome 1 (SRS1)	ID/11p15.5 (180860)	No	5	1
Beckwith–Wiedemann syndrome (BWS)	ID/11p15 ( <i>ICR1</i> , <i>KCNQ1OT1</i> , <i>CDKN1C</i> ) (130650)	No	4	
Silver–Russell syndrome 2 (SRS2)	ID/7p11.2 (180860)	No		
Temple syndrome	ID/14q32 (616222)	No		
Kagami–Ogata syndrome (KOS)	ID/14q32 (608149)	No	2	

OMIM number listed in parentheses adjacent to disorder name. The following list of genes have been classified as having reduced sensitivity and more moderate signatures based on signature strength, limited reference cohort size, or types of variants that have been tested: *CHD8*, *PHF6*, *DNMT3B*, *CDC47*, *ZBTB24*, *HELLS*, *SETD5*, *KMT5B*, *BRWD3*, *ZNF711*, *KAT6B*, *SMS*, *DNMT3A*.

ID imprinting disorder, *TNR* trinucleotide repeat disorder.

<sup>a</sup>*ADNP* has two distinct signatures depending on where in the gene the variant occurs. *HVDAS\_T* signature includes variants that occupy the N- and C-terminus of the gene and *HVDAS\_C* includes variants in the central region of the gene including the nuclear localization signal of the protein.

<sup>b</sup>ICF1 exhibits one signature while ICF 2, 3, and 4 exhibit a separate, common signature.

<sup>c</sup>These two deletion/duplication syndromes exhibit symmetrical increased/decreased DNA methylation signatures, respectively.

tested by EpiSign analysis and reported in the period between September 2019 and June 2020. Of these, 57 specimens were positive for an episignature, 146 were negative, and 4 were inconclusive (Table S3). Of the 207 subjects, a total of 136 patient samples had a previous VUS finding, and of these, 48 (35.3%) had DNA methylation profiles positive for one of the EpiSigns. Most of the positive cases had robust DNA methylation profiles (MVP scores >0.9 with unambiguous MDS and clustering analysis) with some positive cases showing moderate, but positive profiles (reduced but positive MVP score, or closer to borderline MDS clustering). Eighty-six cases had no evidence of a DNA methylation episignature. The remaining two VUS cases were inconclusive.

Figure 1 shows examples of MVP score plots for Cornelia de Lange and Sotos syndromes. Figure S1 shows the representative genomic loci along with reference control and positive DNA methylation profiles for imprinting and fragile X disorders listed in Table 1. For specimens that screen positive for a specific MVP score, hierarchical clustering and multidimensional scaling are performed. Figure 2 shows data for two subjects with VUS in *SMC1A*, the causative gene for Cornelia de Lange syndrome (CdLS) type 2, *SMC1A*:c.598A>C, p.(Lys200Gln) and *SMC1A*:c.1280A>G, p.(Glu427Gly). The *SMC1A*:c.598A>C, p.(Lys200Gln) variant was reclassified as likely pathogenic, ACMG category 2, and the *SMC1A*:c.1280A>G, p.(Glu427Gly) variant did not show evidence of pathogenicity based on the current reference episignatures. Other examples depicted in Fig. 3 show data from two subjects with VUS in *NSD1*, the causative gene for Sotos syndrome type 1, *NSD1*:c.4982G>C, p.(Cys1661Ser) and *NSD1*:c.3331G>T, p.Asp1111Tyr. The *NSD1*:c.4982G>C, p.(Cys1661Ser) was reclassified to likely pathogenic, ACMG category 2, while *NSD1*:c.3331G>T, p.Asp1111Tyr, which clusters with controls, showed no evidence of pathogenicity. Similar analysis is performed for all subjects tested with findings summarized in Table S3.

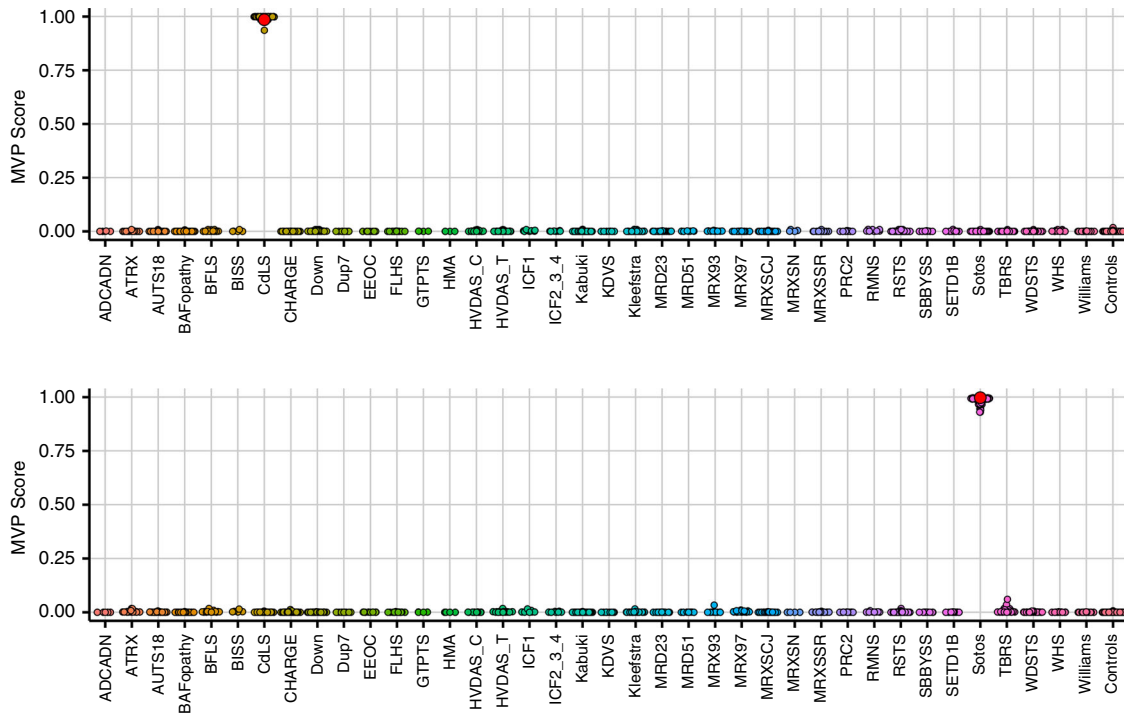
One of the four inconclusive clinical testing cases, Clin77, had a *ARID2*:c.988\_1008del, p.(Leu330\_Gly336del) VUS. MVP score for BAFopathy using the EpiSign v1 was slightly elevated (0.2) but below the established 0.5 cutoff and above the 0.1 cutoff for reference normal controls. The updated EpiSign v2 reanalysis showed no evidence of the elevated BAFopathy score; however, MRD23 score remained elevated. MDS profile showed clustering between BAFopathy cohort and controls (Figure S2). The current

BAFopathy episignature is trained on positive cases with pathogenic variants in *ARID1B*, *ARID1A*, *SMARCB1*, *SMARCA2*, and *SMARCA4*. Since BAF complex-associated *ARID2* positive references are not represented, it was not possible to confidently rule in/out a BAFopathy episignature. Case Clin203 had a previously identified *ADNP*:c.46C>G, p.(Arg16Gly) VUS. The *ADNP* gene has two distinct EpiSigns<sup>11</sup> as a result of truncating variants in two distinct protein domains; the 5' being defined by the variants ranging from c.56 to c.1287. The MVP score for this subject was 0.03, which is within the normal range, but MVP scores for all other conditions were zero. Due to the strong hypomethylation observed with this episignature, MDS analysis clearly separates the reference from the positive cohort, and this sample plots between the two (Figure S2). In lieu of these findings and because the variant lies outside of the established EpiSign domain, the result was reported as inconclusive. For case Clin120, the MVP scores were within the expected reference range but showed slight elevation for MRX97 (0.02) (Figure S2). Although this value is within reference control limits, the currently defined episignature for MRX97 is mild and derived from a limited positive reference cohort. Since this subject's phenotype had a partial overlap with MRX97 the result was reported as inconclusive. The final case, Clin202, was referred because of clinical features consistent with BAFopathy disorder and no variants identified in BAF complex genes. This sample clustered between BAFopathy and the control samples by the MDS analysis (Figure S2). Although the MVP score was within the normal reference range for BAFopathy and all other EpiSign disorders, as in case Clin77, and we could not rule out involvement of other yet unmapped BAF complex genes.

#### Notable clinical cases

While one use of EpiSign is to help resolve VUS, there are scenarios where a DNA methylation episignature is the only molecular diagnostic finding. Case Clin136 was referred for EpiSign analysis due to clinical features consistent with ATRX-related syndrome. However, previous genetic testing of this individual did not identify any alterations in the *ATRX*, using targeted and exome sequencing. In contrast, EpiSign analysis conclusively confirmed presence of a specific *ATRX*-specific DNA methylation signature (Fig. 4). Hence, for this subject, DNA





**Fig. 1 A multiclass supervised classification system.** This classification system, referred to as a methylation variant pathogenicity (MVP) score, can classify samples using the 43 epigenotypes in EpiSign v2. Shown here are examples using the Cornelia de Lange syndrome (CdLS) signature (top) and Sotos signature (bottom), applied to over 1,500 samples from subjects with various neurodevelopmental syndromes or from healthy controls. In each case, samples from the respective syndrome all have high scores while samples from other syndromes and controls all have low scores, demonstrating the sensitivity and specificity of the classifier. The likely pathogenic variants described in Fig. 2 (CdLS) and 3 (Sotos) are shown here larger and in red.

methylation profiling remains the only molecular diagnosis explaining their clinical presentation.

A similar scenario with a different outcome involves subject Clin6 who was referred for EpiSign analysis as a result of negative molecular sequencing (*CHD7* and *SEMA3E*), negative microarray, and negative exome analysis with the phenotype strongly suggestive of CHARGE syndrome. EpiSign analysis identified a DNA methylation profile specific for *CHD7*, consistent with the clinical diagnosis (Fig. 4). As a result, follow-up molecular studies identified the causative deep intronic variant in *CHD7* and confirmed the molecular diagnosis. Details of this case resulting from extensive genomic evaluation by the Undiagnosed Diseases Network (<https://undiagnosed.hms.harvard.edu/>) are described in a separate manuscript in preparation. Hence, EpiSign analysis provided the necessary evidence for a more in-depth and focused molecular analysis, ultimately leading to the definitive molecular genetic diagnosis.

Case Clin187 highlights an example where extensive molecular genetic testing was performed without a molecular diagnosis, referred to as a diagnostic odyssey. Findings included exome sequencing with *PTCHD1:c.605G>A* (p.Arg202Gln) VUS, maternal; *VPS13B:c.1520A>G* (p.Asn507Ser) VUS, paternal; *LAMC3:c.4415G>A* (p.Arg1472Gln), maternal; *EZH2:c.2110+6T>G* VUS, maternal; and normal findings for: *FMR1*, *MECP2*, SNP 6.0 microarray, *NSD1* (sequencing and MLPA), extensive biochemical workup, and X-inactivation studies. However, EpiSign analysis showed an epigenotype associated with *DNMT3A*, the gene involved in Tatton-Brown-Rahman syndrome (Fig. 4). Follow up sequencing of the *DNMT3A* gene revealed a missense likely pathogenic variant *DNMT3A:c.2146G>A* (p.Val716Ile). As an infant, the patient met motor and speech milestones but regressed at 18 months, losing speech, attentiveness, and responsiveness to directions. Now at age 17 years, she has overgrowth (all growth parameters greater

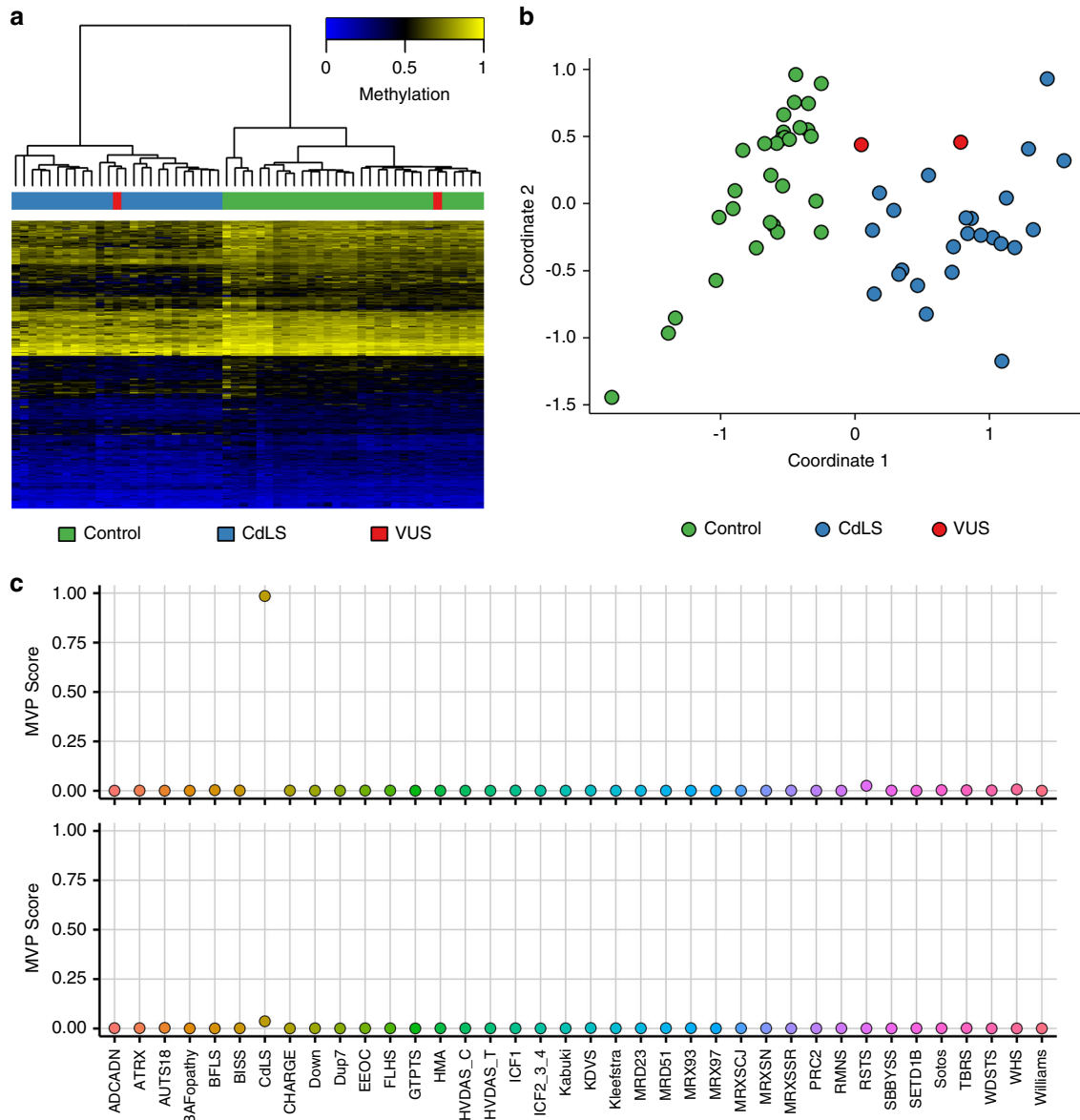
than the 97th centile), limited speech, echolalia, mild intellectual disability, and autism spectrum disorder.

## DISCUSSION

This paper describes the implementation of genome-wide DNA methylation analysis in clinical testing of individuals with rare genetic disorders. It represents a major milestone in molecular diagnostics as it advances clinical genetic testing beyond assessment of the DNA nucleotide sequence, or genomics, to clinical epigenomics. While targeted molecular assays for assessment of DNA methylation defects of imprinting disorders<sup>28,29</sup> and fragile X syndrome<sup>30</sup> have been around for decades, EpiSign enables simultaneous assessment for these, as well as a rapidly expanding number of genetic neurodevelopmental disorders exhibiting DNA methylation epigenotypes.

### Clinical impact

The primary clinical utility of EpiSign analysis is the assessment and reclassification of VUS in genes with existing epigenotypes,<sup>15</sup> and the assessment of genetically unsolved individuals with suspected hereditary conditions.<sup>14</sup> While this clinical utility has been described in a research setting in targeted patient cohorts, this study focuses on the assessment of the impact of genome-wide DNA methylation analysis in a clinical setting in a prospective unselected patient population. The majority of the subjects tested in the clinical cohort (136; 65.7%) had previously identified genetic VUS, and of these, 134 (98.5%) received a conclusive EpiSign report (positive or negative), of which 48 (35.3%) showed a distinct EpiSign. For the 86 (63.2%) VUS cases with negative EpiSign results, lack of an epigenotype is considered strong evidence for rule out of these disorders. However, due to the possibility of the existence of additional/alternate yet unmapped

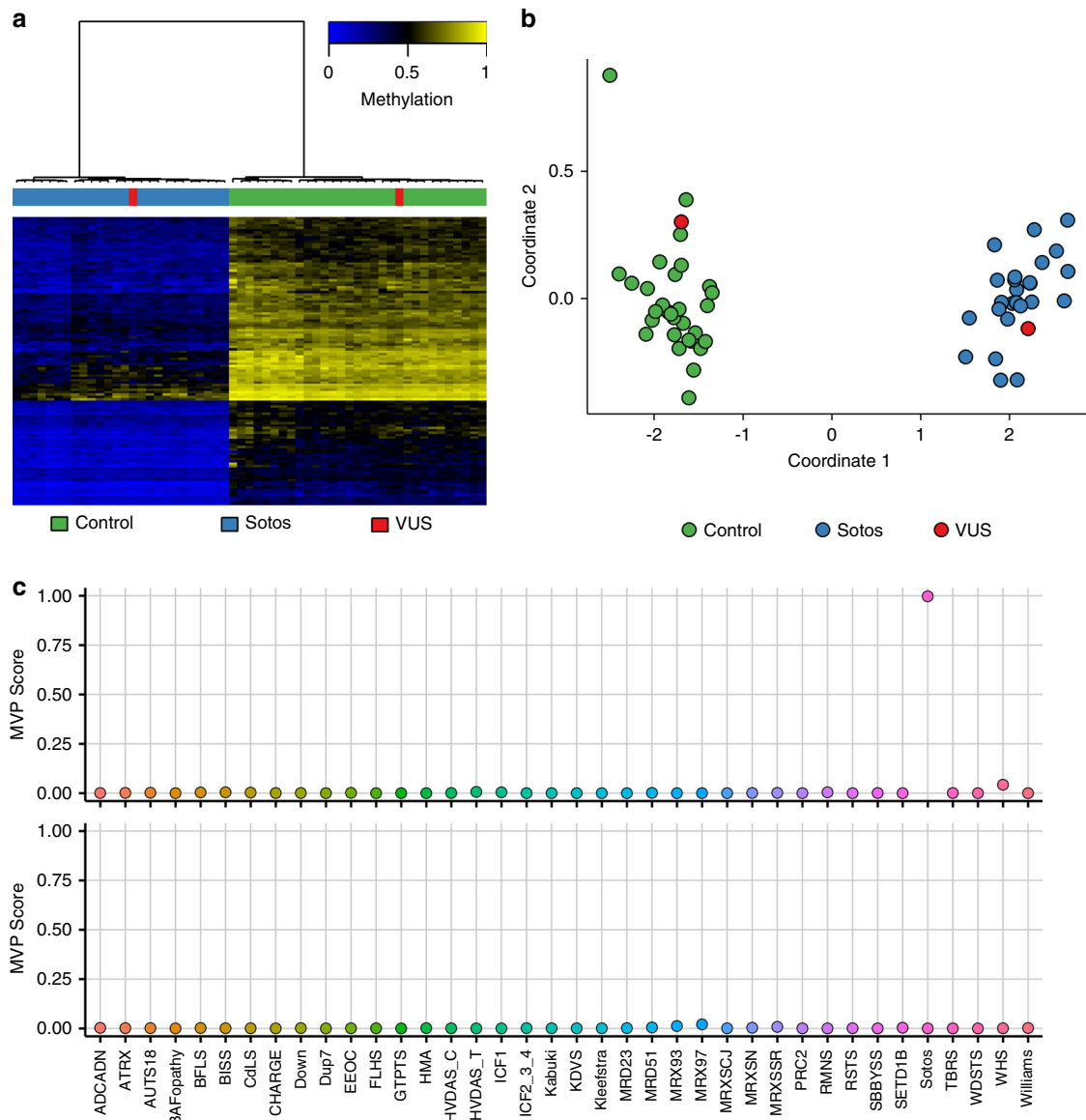


**Fig. 2 DNA methylation (EpiSign) analysis of peripheral blood from two subjects with variants of unknown clinical significance (VUS) in *SMC1A*, the causative gene for Cornelia de Lange syndrome (CdLS) type 2.** The variants are *SMC1A*:c.598A>C, p.(Lys200Gln) (labeled red and clustering with *CdLS* samples) and *SMC1A*:c.1280A>G, p.(Glu427Gly) (labeled red and clustering with control samples). (a) Hierarchical clustering and (b) multidimensional scaling plot of subjects with a confirmed CdLS epigenature, controls, and the VUS under investigation. The *SMC1A*:c.598A>C VUS clustering with the CdLS samples indicates it has a DNA methylation signature similar to that seen in other CdLS samples and suggesting that the variant is pathogenic, while *SMC1A*:c.1280A>G is likely benign. (c) Methylation variant pathogenicity (MVP) score, a multiclass supervised classification system capable of discerning between the 43 different epigenatures in EpiSign V2, was applied to the *SMC1A*:c.598A>C likely pathogenic variant (top) and the *SMC1A*:c.1280A>G likely benign variant (bottom). This classification system generates a probability score for each epigenature, with a score near 1 indicating that the sample has an epigenature similar to the reference epigenature.

epigenatures in these genes, as has been demonstrated in *ADNP*<sup>11</sup> and *SMARCA2*,<sup>31</sup> we currently stop short of fully ruling out pathogenicity. Only 2 (1.5%) VUS cases remained unclassified. These findings represent a significant advancement in clinical variant assessment over currently available methodologies. While alternative approaches for classification of VUS exist, including functional and family segregation studies, these are not always available, feasible, or conclusive.<sup>32</sup> EpiSign analysis is less susceptible to those limitations as it assesses the same tissue (patient's peripheral DNA) used for sequence analysis, and it generally does not require assessment of other family members. From a patient, family, and clinical management perspective,

providing the patient and care team with a rapid diagnosis relating to the functional impact of the genetic condition, which for a microarray-based test can be achieved within 1–2 weeks, may be most beneficial to decision-making.

Of the remaining 71 subjects without the previously identified genetic findings, 8 subjects (11.3%) had a positive EpiSign result, demonstrating evidence for the clinical utility of EpiSign analysis in a broader patient population. Given that these subjects already had varying degrees of targeted genetic investigations completed with no conclusive findings, some of the EpiSign disorders had effectively been ruled out already. These findings taken together suggest there may be a health systems value proposition of



**Fig. 3** DNA methylation (EpiSign) analysis of peripheral blood from two subjects with variants of unknown clinical significance (VUS) in *NSD1*, the causative gene for Sotos syndrome type 1. The variants are *NSD1*:c.4982G>C,p.Cys1661Ser (labeled red and clustering with *NSD1* samples) and *NSD1*:c.3331G>T,p.Asp1111Tyr (labeled red and clustering with control samples). (a) Hierarchical clustering and (b) multidimensional scaling plot of subjects with a confirmed Sotos epigenature, controls, and the VUS under investigation. The *NSD1*:c.4982G>C VUS clustering with the Sotos samples indicates it has a DNA methylation signature similar to that seen in other Sotos samples and suggesting that the variant is pathogenic, while *NSD1*:c.3331G>T is likely benign. (c) Methylation variant pathogenicity (MVP) score, a multiclass supervised classification system capable of discerning between the 43 different epigenatures in EpiSign v2, was applied to the *NSD1*:c.4982G>C likely pathogenic variant (top) and the *NSD1*:c.3331G>T likely benign variant (bottom). This classification system generates a probability score for each epigenature, with a score near 1 indicating that the sample has an epigenature similar to the reference epigenature.

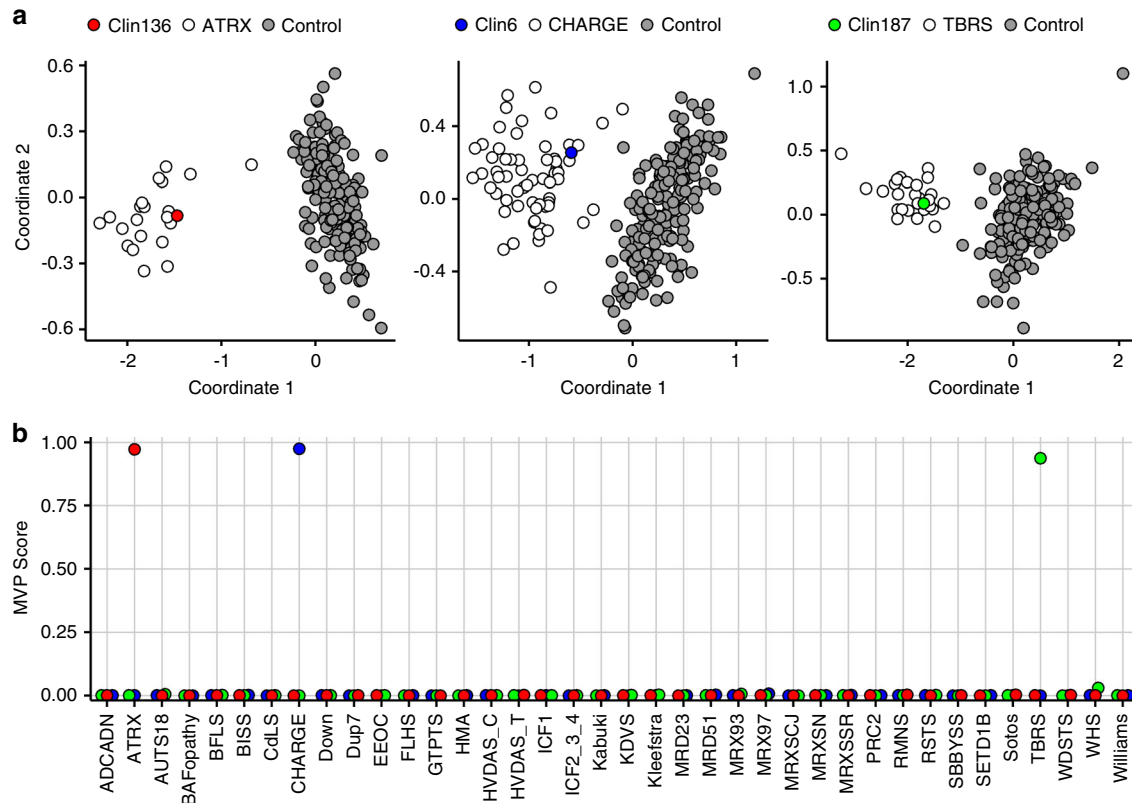
instigating use of EpiSign earlier in the diagnostic journey of individuals with rare disorders.

Application of this technology to the broader patient populations will depend on the rate of discovery of gene and disorder-specific epigenatures. As a corollary, implementation of chromosomal microarrays as a first-tier diagnostic test was primarily contingent upon increased diagnostic yield compared with karyotyping (from 5% to 10–15%) resulting from years of research and discovery of novel microdeletion and duplication syndromes, often involving large clinical databases and registries.<sup>33,34</sup> Similarly, there are now major efforts underway to assess the clinical utility and the health systems impact, and to accelerate the rate of epigenature discovery including a national-scale trial

(“Beyond Genomics: Assessing the Improvement in Diagnosis of Rare Diseases using Clinical Epigenomics in Canada [EpiSign-CAN]”), which will compare the impact of DNA methylation analysis as a first-line versus a second-line test in 4,000 individuals with suspected rare disorders while assessing EpiSigns in 100 additional genetic conditions (<https://www.genomecanada.ca/en/beyond-genomics-assessing-improvement-diagnosis-rare-diseases-using-clinical-epigenomics-canada>).

The clinical cases presented in detail highlight some important implications of this technology. One is the sequential use of EpiSign analysis with genetic testing. While using EpiSign to investigate individuals with VUS or related clinical presentations provides demonstrated value, EpiSign can also uncover genetic





**Fig. 4 Clinical case EpiSign assignment.** Assessment of the notable clinical cases. (a) Principal component analysis (PCA) plots for Clin136 (no genetic variant identified), Clin6 (no genetic variant identified), and Clin187 (*DNMT3A:c.2146G>A* [p.Val716Ile]). (b) Corresponding methylation variant pathogenicity (MVP) scores for the notable clinical cases. Both PCA plots and MVP scores provide conclusive evidence for EpiSign classification of these clinical cases.

disorders that were not initially suspected. We have previously demonstrated an incremental diagnostic yield of approximately 3% in patients with prior extensive genomic testing but without a genetic diagnosis.<sup>9,14</sup> Existence of a specific DNA methylation pattern can guide the molecular assessment, and in some cases resolve complex diagnostic odysseys, which can have a huge impact on patient care and the related health systems costs.<sup>35,36</sup>

#### Clinical service delivery

EpiSign testing is performed using an integrated model involving primary labs performing and reporting test results with informatics and databasing centralized in the tertiary clinical bioinformatics laboratory, similar to the field of noninvasive prenatal testing for aneuploidy.<sup>37</sup> The key benefits of this model involve standardization and coordinated quality management and quality assessment procedures, ensuring consistency across the different provider laboratories, which is critical given the inherent nature and complexity of this analysis. This also allows for more rapid expansion of reference DNA methylation databases and facilitates continuous optimization of the underpinning analytical algorithms. As the EKD expands, the reference machine learning-derived algorithms that form the basis of individual EpiSigns become more sensitive and specific,<sup>9</sup> enabling regular and documented updates to the analytical software, with the appropriate quality metrics and quality control documentation and version controls.

#### Limitations

There are a number of challenges related to introducing a diagnostic modality to the clinical laboratory, and EpiSign is not an exception. Unlike DNA sequencing, DNA methylation analysis is

limited to peripheral blood where large reference databases are available. Other factors including age, sex, and environmental exposures can also impact the analysis and need to be accounted for in analytical processes.<sup>38,39</sup>

DNA methylation epigenatures can be susceptible to technical variation such as sample processing data batch effects, as well as biological parameters such as mosaicism. The validation cohort included a number of low-level mosaic imprinting disorders that may not be readily detectable by the EpiSign assay. In our previous work we were able to detect mosaicism in imprinting disorders<sup>24</sup> and fragile X,<sup>25</sup> at levels >20%, but due to normal control variability (Figure S1), this is currently not routinely possible for samples with lower-level mosaicism. Mosaicism is also a limitation for detection of other EpiSigns, and may provide an explanation for some of the discordant samples. We have previously demonstrated that total gene dosage dilutes the intensity of the EpiSign signal, as in heterozygous females in the X-linked *KDM5C*-related Claes–Jensen syndrome;<sup>16</sup> however, here, as a result of having a reference cohort we were able to derive a specific and sensitive MVP score for heterozygous females. In addition to mosaicism, a possible reason for nonconcordance in a laboratory setting could be sample mix up. Alternatively, discordance of sample Val26 with *CREBBP:c.4480C>A*, p.(Pro1494Thr) may be explained by the variant not actually being pathogenic, or an existence of a yet unmapped epigenature in this gene.

Other biologically based limitations of EpiSign analysis can be highlighted by the four inconclusive samples from the clinical testing cohort. The current BAFopathy epigenature is trained on positive cases with pathogenic variants in *ARID1B*, *ARID1A*, *SMARCB1*, *SMARCA2*, and *SMARCA4*.<sup>17</sup> Hence, it is not possible to completely rule out pathogenicity of a variant of another BAF

complex gene as for the case with the *ARID2* variant. An alternative scenario is highlighted by *ADNP*, which was the first gene in which the existence of multiple, domain-specific epesignatures was described.<sup>11</sup> Although majority of the EpiSign genes currently have a single common epesignature mapped, assessment of pathogenicity of variants outside the established reference range and variant type warrants caution, as in the inconclusive case Clin203.

An overarching challenge with this technology is the rarity of Mendelian disorders. While the population prevalence of rare diseases is 3.5–5.9%, equating to 263–446 million persons affected globally, given that this number encompasses >5,000 diseases, the prevalence of rare disorders ranges between 1–5 per 10,000 and <1/1,000, 000.<sup>40</sup> Generation of EpiSigns requires cohorts of subjects with gene-specific pathogenic variants, which is currently possible for the more prevalent disorders. Also, as many of the epesignatures are mild in scale, the size of the reference cohort is directly correlated to the level of sensitivity of the assay. Hence, occasionally, as in the case Clin120 for example, the results may be inconclusive.

While the recommendations for application of the functional evidence in genetic testing now exist,<sup>6</sup> there are currently no specific guidelines for the clinical interpretation of genomic DNA methylation findings. EpiSign employs the use of DNA methylation data as a surrogate for evidence of the underpinning genetic defects that may or may not be detectable using the current molecular testing modalities. Another challenge is that our current knowledge of the full scope of these genetically associated epiphenotypes is limited. Hence, while on the one hand a confirmation of a related epesignature in a patient with a genetic VUS may be considered a molecular diagnosis, a negative result in a patient with or without a known genetic variant is not an absolute rule-out and the analysis would be considered a molecular screen. A more complex challenge is a confirmation of an epesignature in absence of a detectable genetic variation, in particular when it is observed in a patient with a matching clinical diagnosis. In this case EpiSign is the standalone molecular diagnosis.

## Conclusion

This study describes the implementation of clinical genomic DNA methylation testing in patients with rare disorders. It demonstrates strong evidence of clinical utility, including the ability to provide conclusive diagnoses in a significant proportion of subjects tested. It also highlights the limitations and challenges with implementation and use of this diagnostic modality. As this technology evolves, the number and type of rare disorders with EpiSigns is going to expand, increasing its clinical utility. While the current clinical use of EpiSign focuses on cases with VUS, and very specific clinical presentations, expanding clinical utility of this test may justify its application earlier in the diagnostic journey, in a broader patient population. Larger-scale studies, such as EpiSign-CAN, are necessary to assess the diagnostic yield and health system impact as either a first-line test or in unresolved cases post-genomic assessment. Finally, the development of clinical guidelines for use and application of clinical epigenomic technologies is warranted.

## DATA AVAILABILITY

The summarized, anonymized data for each subject are described in the study. The raw anonymized DNA methylation data are available from the authors upon request. Software used in this study is publicly available and detailed analytical methodology is as previously reported.<sup>9</sup>

Received: 27 August 2020; Revised: 29 December 2020; Accepted: 31 December 2020

Published online: 5 February 2021

## REFERENCES

- Christianson, A., Howson, C. & Modell, B. Global report on birth defect. The hidden toll of dying and disabled children. (New York, March of Dimes, 2006).
- Baird, P. A., Anderson, T. W., Newcombe, H. B. & Lowry, R. B. Genetic disorders in children and young adults: a population study. *Am J Hum Genet* **42**, 677–693 (1988).
- Kvarnang, M. & Nordgren, A. Intellectual disability & rare disorders: a diagnostic challenge. *Adv. Exp. Med. Biol.* [https://doi.org/10.1007/978-3-319-67144-4\\_3](https://doi.org/10.1007/978-3-319-67144-4_3) (2017).
- Schwarze, K., Buchanan, J., Taylor, J. C. & Wordsworth, S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet. Med.* **20**, 1122–1130 (2018).
- Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* <https://doi.org/10.1038/gim.2015.30> (2015).
- Brnich, S. E. et al. Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med.* 2019. <https://doi.org/10.1186/s13073-019-0690-2> (2019).
- Schenkel, L. C., Rodenhiser, D. I., Ainsworth, P. J., Paré, G. & Sadikovic, B. DNA methylation analysis in constitutional disorders: clinical implications of the epigenome DNA methylation analysis in constitutional disorders: clinical implications of the epigenome. *Crit Rev Clin Lab Sci.* <https://doi.org/10.3109/10408363.2015.1113496> (2016).
- Schübeler, D. Function and information content of DNA methylation. *Nature.* <https://doi.org/10.1038/nature14192> (2015).
- Aref-Eshghi, E. et al. Evaluation of DNA methylation epesignatures for diagnosis and phenotype correlations in 42 Mendelian neurodevelopmental disorders. *Am. J. Hum. Genet.* <https://doi.org/10.1016/j.ajhg.2020.01.019> (2020).
- Cioffi, A. et al. Frameshift mutations at the C-terminus of HIST1H1E result in a specific DNA hypomethylation signature. *Clin Epigenetics.* <https://doi.org/10.1186/s13148-019-0804-0> (2020).
- Bend, E. G. et al. Gene domain-specific DNA methylation epesignatures highlight distinct molecular entities of ADNP syndrome. *Clin Epigenetics.* <https://doi.org/10.1186/s13148-019-0658-5> (2019).
- Krzyzewska, I. M. et al. A genome-wide DNA methylation signature for SETD1B-related syndrome. *Clin Epigenetics.* <https://doi.org/10.1186/s13148-019-0749-3> (2019).
- Aref-Eshghi, E. et al. Genome-wide DNA methylation and RNA analyses enable reclassification of two variants of uncertain significance in a patient with clinical Kabuki syndrome. *Hum. Mutat.* <https://doi.org/10.1002/humu.23833> (2019).
- Aref-Eshghi, E. et al. Diagnostic utility of genome-wide DNA methylation testing in genetically unsolved individuals with suspected hereditary conditions. *Am. J. Hum. Genet.* <https://doi.org/10.1016/j.ajhg.2019.03.008> (2019).
- Aref-Eshghi, E. et al. Genomic DNA methylation signatures enable concurrent diagnosis and clinical genetic variant classification in neurodevelopmental syndromes. *Am. J. Hum. Genet.* <https://doi.org/10.1016/j.ajhg.2017.12.008> (2018).
- Schenkel, L. C. et al. Peripheral blood epi-signature of Claes-Jensen syndrome enables sensitive and specific identification of patients and healthy carriers with pathogenic mutations in KDM5C. *Clin Epigenetics.* <https://doi.org/10.1186/s13148-018-0453-8> (2018).
- Aref-Eshghi, E. et al. BAFopathies' DNA methylation epi-signatures demonstrate diagnostic utility and functional continuum of Coffin–Siris and Nicolaides–Baraitser syndromes. *Nat. Commun.* <https://doi.org/10.1038/s41467-018-07193-y> (2018).
- Kernohan, K. D. et al. Identification of a methylation profile for DNMT1-associated autosomal dominant cerebellar ataxia, deafness, and narcolepsy. *Clin Epigenetics.* <https://doi.org/10.1186/s13148-016-0254-x> (2016).
- Hood, R. L. et al. The defining DNA methylation signature of Floating Harbor syndrome. *Sci. Rep.* <https://doi.org/10.1038/srep38803> (2016).
- Butcher, D. T. et al. CHARGE and Kabuki syndromes: gene-specific DNA methylation signatures identify epigenetic mechanisms linking these clinically overlapping conditions. *Am. J. Hum. Genet.* <https://doi.org/10.1016/j.ajhg.2017.04.004> (2017).
- Choufani, S. et al. NSD1 mutations generate a genome-wide DNA methylation signature. *Nat. Commun.* <https://doi.org/10.1038/ncomms10207> (2015).
- Choufani, S. et al. DNA methylation signature for EZH2 functionally classifies sequence variants in three PRC2 complex genes. *Am. J. Hum. Genet.* <https://doi.org/10.1016/j.ajhg.2020.03.008> (2020).
- Sadikovic, B., Aref-Eshghi, E., Levy, M. A. & Rodenhiser, D. DNA methylation signatures in mendelian developmental disorders as a diagnostic bridge between genotype and phenotype. *Epigenomics.* <https://doi.org/10.2217/epi-2018-0192> (2019).

24. Aref-Eshghi, E. et al. Clinical validation of a genome-wide DNA methylation assay for molecular diagnosis of imprinting disorders. *J. Mol. Diagn.* <https://doi.org/10.1016/j.jmoldx.2017.07.002> (2017).
25. Schenkel, L. C. et al. Clinical validation of fragile X syndrome screening by DNA methylation array. *J. Mol. Diagn.* <https://doi.org/10.1016/j.jmoldx.2016.06.005> (2016).
26. Platt, J. et al. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Adv Large Margin Classif.* **10** (2000).
27. Sadikovic, B., Levy, M. A. & Aref-Eshghi, E. Functional annotation of genomic variation: DNA methylation epigenatures in neurodevelopmental Mendelian disorders. *Hum. Mol. Genet.* <https://doi.org/10.1093/hmg/ddaa144> (2020).
28. Li, E., Beard, C. & Jaenisch, R. Role for DNA methylation in genomic imprinting. *Nature.* <https://doi.org/10.1038/366362a0> (1993).
29. Sutcliffe, J. S. et al. Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. *Nat Genet.* <https://doi.org/10.1038/ng0994-52> (1994).
30. Sutcliffe, J. S. et al. DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum. Mol. Genet.* <https://doi.org/10.1093/hmg/1.6.397> (1992).
31. Cappuccio, G. et al. De novo SMARCA2 variants clustered outside the helicase domain cause a new recognizable syndrome with intellectual disability and blepharophimosis distinct from Nicolaides-Baraitser syndrome. *Genet. Med.* <https://doi.org/10.1038/s41436-020-0898-y> (2020).
32. Cooper, G. M. Parlez-vous VUS? *Genome Res.* <https://doi.org/10.1101/gr.190116.115> (2015).
33. Miller, D. T. et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* <https://doi.org/10.1016/j.ajhg.2010.04.006> (2010).
34. South, S. T., Lee, C., Lamb, A. N., Higgins, A. W. & Kearney, H. M. ACMG standards and guidelines for constitutional cytogenomic microarray analysis, including postnatal and prenatal applications: revision 2013. *Genet. Med.* <https://doi.org/10.1038/gim.2013.129> (2013).
35. Lazaridis, K. N. et al. Outcome of whole exome sequencing for diagnostic odyssey cases of an individualized medicine clinic: the Mayo Clinic Experience. *Mayo Clin. Proc.* <https://doi.org/10.1016/j.mayocp.2015.12.018> (2016).
36. Srivastava, S. et al. Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genet. Med.* <https://doi.org/10.1038/s41436-019-0554-6> (2019).
37. Benn, P., Cuckle, H. & Pergament, E. Non-invasive prenatal testing for aneuploidy: current status and future prospects. *Ultrasound Obstet. Gynecol.* <https://doi.org/10.1002/uog.12513> (2013).
38. Christensen, B. C. et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CPG island context. *PLoS Genet.* <https://doi.org/10.1371/journal.pgen.1000602> (2009).
39. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* <https://doi.org/10.1186/gb-2013-14-10-r115> (2013).
40. Nguengang Wakap, S. et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur. J. Hum. Genet.* <https://doi.org/10.1038/s41431-019-0508-0> (2020).

## ACKNOWLEDGEMENTS

Funding for this study is provided in part by the London Health Sciences Molecular Diagnostics Development Fund and Genome Canada Genomic Applications Partnership Program.

## ETHICS DECLARATION

Participants provided informed consent for clinical testing prior to sample collection and EpiSign analysis. All of the samples and records were de-identified. Informed consent was obtained by the physicians to use specific clinical information for the three cases that were described in the "Notable Clinical Cases" section. The research was conducted in accordance with all relevant ethical regulations. The study protocol has been approved by the Western University Research Ethics Board (REB 106302).

## 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/s41436-020-01096-4>

**Correspondence** and requests for materials should be addressed to B.S., M.L.T. or M.A.

**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.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021, corrected publication 2021