

# CpG methylation signature defines human temporal lobe epilepsy and predicts drug-resistant

Wenbiao Xiao<sup>1</sup>  | Chaorong Liu<sup>1</sup> | Kuo Zhong<sup>2</sup> | Shangwei Ning<sup>3</sup> | Rui Hou<sup>4</sup> | Na Deng<sup>1</sup> | Yuchen Xu<sup>1</sup> | Zhaohui Luo<sup>1</sup> | Yujiao Fu<sup>1</sup> | Yi Zeng<sup>5</sup> | Bo Xiao<sup>1</sup>  | Hongyu Long<sup>1</sup>  | Lili Long<sup>1</sup>

<sup>1</sup>Department of Neurology, Xiangya Hospital, Central South University, Changsha, China

<sup>2</sup>Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen, China

<sup>3</sup>College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China

<sup>4</sup>Shanghai Biotechnology Corporation, Shanghai, China

<sup>5</sup>Department of Geriatrics, Second Xiangya Hospital, Central South University, Changsha, China

## Correspondence

Hongyu Long and Lili Long, Department of Neurology, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China. Emails: longhongyu@csu.edu.cn (HL); 10353654@qq.com (LL)

## Funding information

Fundamental Research Funds for the Central Universities of Central South University, Grant/Award Number: 2018zts248; Clinical Research Foundation of Xiangya Hospital, Grant/Award Number: 2016L08; Hunan Provincial Science and Technology Department, Grant/Award Number: 2019SK1012; National Natural Science Foundation of China, Grant/Award Number: 81671299, 81671300, 81701182 and 81974206

## Abstract

**Aims:** Temporal lobe epilepsy (TLE) is the most common focal epilepsy syndrome in adults and frequently develops drug resistance. Studies have investigated the value of peripheral DNA methylation signature as molecular biomarker for diagnosis or prognosis. We aimed to explore methylation biomarkers for TLE diagnosis and pharmacoresistance prediction.

**Methods:** We initially conducted genome-wide DNA methylation profiling in TLE patients, and then selected candidate CpGs in training cohort and validated in another independent cohort by employing machine learning algorithms. Furthermore, nomogram comprising DNA methylation and clinicopathological data was generated to predict the drug response in the entire patient cohort. Lastly, bioinformatics analysis for CpG-associated genes was performed using Ingenuity Pathway Analysis.

**Results:** After screening and validation, eight CpGs were identified for diagnostic biomarker with an area under the curve (AUC) of 0.81 and six CpGs for drug-resistant prediction biomarker with an AUC of 0.79. The nomogram for drug-resistant prediction comprised methylation risk score, disease course, seizure frequency, and hippocampal sclerosis, with AUC as high as 0.96. Bioinformatics analysis indicated drug response-related CpGs corresponding genes closely related to DNA methylation.

**Conclusions:** This study demonstrates the ability to use peripheral DNA methylation signature as molecular biomarker for epilepsy diagnosis and drug-resistant prediction.

## KEYWORDS

biomarker, DNA methylation, machine learning, nomogram, temporal lobe epilepsy

Xiao and Liu are contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *CNS Neuroscience & Therapeutics* published by John Wiley & Sons Ltd

## 1 | INTRODUCTION

Temporal lobe epilepsy (TLE) is the most common focal epilepsy syndrome in adults and frequently develops drug resistance,<sup>1</sup> requiring surgical treatment which offers a comparatively favorable prognosis.<sup>2,3</sup> Moreover, cognitive impairment and psychiatric comorbidities including depression and anxiety disorders, together with the long-term actual seizures and accompanying drug usage, often result in severe effects on the quality of life and individual health.<sup>4,5</sup>

At present, the diagnosis of epilepsy mainly depends on clinical manifestation, neuroimaging, and electroencephalogram (EEG). These methods are not only expensive and time-consuming, but also require professional equipment and trained specialists that are not accessible to many patients, which result in delayed diagnosis or misdiagnosis to some extent.<sup>6,7</sup> Furthermore, drug-response prediction is mainly based on subjective clinical features by experience and has not come to a conclude.<sup>4,8,9</sup> Earlier identification of drug-resistant patients makes it possible to benefit from epilepsy surgery. Thus, biomarkers for assisting the current diagnosis and predicting the treatment outcome are in urgent need. Preliminary attempts have been made in circulating molecules biomarkers of epilepsy, including inflammatory cytokines, S100 calcium-binding protein B (S100B), and matrix metalloproteinase 9 (MMP9), and recently miRNA.<sup>10-14</sup> However, the limitations of these studies mainly related to small sample size and lack of validation, as well as heterogeneity of epilepsy that prevent the clinical value of these biomarkers.<sup>15</sup>

DNA methylation, the best-studied epigenetic mechanism, refers to the covalent attachment of methyl groups to the cytosine residues (mainly confined in CpG sites) mediated by DNA methyltransferase (DNMT).<sup>16,17</sup> It is mostly stable throughout the genome and is associated with transcriptional activation/repression.<sup>18,19</sup> Aberrant DNA methylation implicated in underlying epileptogenesis and progression mechanisms of epilepsy has gained considerable attention. Altered expression of DNMTs and methylation changes in individual candidate genes (ie, *RELN*) have been found in TLE patients.<sup>20-23</sup> Several genome-wide studies using epileptic brain tissue have identified differential methylation events occurred in genes associated with inflammation, neuronal development, etc.<sup>24-27</sup> Moreover, our previous research reported that dysregulated methylation implicated in both protein-encoding genes and noncoding RNA genes in peripheral blood DNA from TLE patients.<sup>28,29</sup>

A substantial number of studies have investigated the value of peripheral DNA methylation signature as molecular biomarker for diagnosis or prognosis, especially in cancer research.<sup>15,30-32</sup> The prognostic value of O<sup>6</sup>-methylguanine-DNA-methyltransferase (*MGMT*) promoter methylation in glioblastoma and methylated *SEPTIN 9* (*SEPT9*) in plasma for detection of asymptomatic colorectal cancer is well-known paradigms,<sup>33-35</sup> which have been included in clinical guidelines and translated into the commercially available clinical test.<sup>15</sup> In addition, there was a trend that researchers favored combinatorial biomarkers of multiple CpG signature.<sup>36-39</sup> DNA methylation-based biomarkers present advantages with regard to clinical application: presence in various biofluids, more stable than

other biological materials (such as RNA or protein), easy detection by well-established methodologies, and cell-type specificity.<sup>15,31</sup> However, to date, methylation biomarkers for TLE diagnosis and pharmacoresistance prediction have not been explored.

In this study, we aimed to identify and validate disease-related and drug response-related CpGs in TLE. We initially conducted genome-wide DNA methylation profiling in TLE patients; then, we selected candidate CpGs in training cohort and validated those CpGs in another independent cohort by employing machine learning algorithms. Furthermore, a nomogram comprising DNA methylation and clinicopathological data was generated to predict the drug response in the entire patient cohort. Lastly, mechanistic links were pursued for all biomarker CpGs corresponding genes by bioinformatics analysis.

## 2 | MATERIALS AND METHODS

### 2.1 | Patient cohorts

The study was carried out on a cohort of 78 patients with TLE and 78 sex- and age-matched healthy controls, from the Department of Neurology at Xiangya Hospital. And all patients went through comprehensive medical history, physical examination, cranial magnetic resonance imaging (MRI) scans, and EEG. Inclusion criteria of TLE and drug-resistant epilepsy were accorded to our previous research.<sup>28</sup> Written informed consent was obtained from all enrolled participants. Study was conducted in accordance with the guideline for the research involving human and approved by the Ethics Committee of Central South University, Xiangya School of Medicine and the affiliated Xiangya Hospital (201303120). The data were divided into two sets: in the training cohort, 30 TLE patients were analyzed; in the validation phase, candidate CpGs were validated in another independent cohort (n = 48).

### 2.2 | DNA methylation quality control and processing

Whole blood DNA extraction and quality control were constructed as in our previous study.<sup>28</sup> The discovery and training samples were run on the Illumina Infinium HumanMethylation450 BeadChip Kit (450K array). The validation samples were run on the Illumina Infinium HumanMethylationEPIC BeadChip (850K array). The samples DNA underwent bisulfite treatment using the EZ-96 DNA Methylation kit (Zymo Research Corporation, Irvine, CA, USA) and hybridized to arrays according to Illumina recommended protocols. All samples passed the Illumina quality control. Methylation at individual CpG was reported as a methylation  $\beta$ -value, ranging continuously from 0 (unmethylated) to 1 (completely methylated). The minfi R package (Version 1.18.1) was used to retrieve raw data of 450K and 850K array. Initially, we excluded probes located on the sex chromosome and null probes. We also removed the failed probes with a

detection  $P$ -value  $> .05$  in more than 5% samples. The probes with single-nucleotide polymorphisms of MAF  $> 5\%$  within 10 bp of the CpG sites were also rejected. We next performed Subset-quantile Within Array Normalization (SWAN) methods for normalization.<sup>40</sup> The probes of 450K assay are expected to perform similarly on data from the 850K array. In this study, we removed the set of CpG sites that were not included in the 850K array.

### 2.3 | Building a diagnostic model

We included three phases to identify and validate disease-related CpGs signature for patients with TLE. In the discovery phase, the logical regression test was performed to obtain differentially methylated CpG sites (DMCs) between 30 TLE and 30 normal control samples, with a threshold value of .001 for  $P$ -value was used subsequently for filtration. The 237 candidate CpGs were analyzed by Least Absolute Shrinkage and Selection Operator (LASSO) methods. The CpGs were then ranked by the regression parameters. In the training phase, to further shrink the marker numbers to a reasonable range, support vector machine (SVM) algorithm was used for different number of CpGs. As a result, eight CpGs with the highest prediction accuracy were confirmed. SVM algorithms were tuned by 5-fold internal cross-validation, which implies optimal determination of parameters of the SVM algorithm. In the validation phase, the parameters of the SVM model from the training cohort were used to an independent cohort of 96 samples (48 TLE and 48 normal controls) for validating the diagnostic performance of the model.

### 2.4 | Building a predictive model for drug response

In the discovery phase, the  $t$  test was performed to identify DMCs between 10 drug-resistant and 20 drug-responsive samples, with a threshold value of .005 for  $P$ -value. After 99 DMCs were obtained, we used SVM-Recursive Feature Elimination (SVM-RFE) to select candidate CpG sites. In the training phase, logistic regression was used to further narrow CpGs. Six CpGs with the highest prediction accuracy were identified, with parameter tuning conducted by 5-fold cross-validation. A risk score was calculated for each patient using a formula derived from the methylation levels of these 6 CpGs weighted by their regression coefficient. Validation analyses were performed in another cohort (17 drug-resistant and 13 drug-responsive samples). In addition, a nomogram comprising integrated DNA methylation risk score and clinicopathological data was generated to predict the drug response. The performance of the nomogram was explored graphically by calibration plots.

### 2.5 | Bioinformatics analysis

Pathway analysis for CpG-associated genes was performed using Ingenuity Pathway Analysis (IPA; <http://www.ingenuity.com/>). For

the purposes of this study, the canonical pathway and diseases functions analysis available in IPA were applied, which resulted in the inclusion of CpG corresponding genes and the other identified genes interacting with in the analysis. The Fisher's exact test was applied to measure the significance of the association between genes mapped by IPA and the canonical pathway.

### 2.6 | Bisulfite pyrosequencing of selected DNA methylation loci

Bisulfite pyrosequencing is well-established technique that used for quantitative methylation analysis of genomic regions in single-nucleotide resolution.<sup>41</sup> We selected 4 CpG loci (cg25838818, cg27564766, cg11954680, and cg26119877) for assay cross-validation by bisulfite pyrosequencing. Blood DNA samples from 10 TLE patients and 10 healthy control cases or 10 drug-responsive TLE cases and 10 drug-resistant TLE cases were bisulfite converted, followed by PCR amplification of the relevant regions using the PyroMark PCR kit (Qiagen, CA, USA) according manufacturer's instruction. Nucleotide probes with biotinylated version can be detected by streptavidin sepharose, as listed below:

(a) cg25838818: GTAGTTGAGGGTTAGGAAAGATGTG (F), ATACA AATACCAACTCCCTCTAATTCAT (R) and GTGAAAAATTTTGTG TGGTG (S).

(b) cg27564766: GGAGGGATAGGGTTGTTT (F), CCAACCAA CCACCTCATC (R) and TTGTGGTGGTTTATAGG (S).

(c) cg11954680: ATTAGAATTAAGAGTGATTTAGGAAGTG (F), AA AAAAAATTTTCTATTTCACCTTCTA (R) and GTGATTTAGGAAG TGGTTAA (S).

(d) cg26119877: AAGATTGGGTGGTTTATAAGAAAG (F), CCACA AATAAACACATTTTACTATAACAC (R) and GTTGTGGGATTAG TTTG (S).

Pyrosequencing assay, purification, and subsequent processing of the biotinylated single-stranded DNA were carried out according to the manufacturer's recommendations.

### 2.7 | Statistical analysis

In the comparative analysis of clinical characteristics (SPSS18.0), measurement data (age, disease course, and seizure frequency) were subject to  $K$ - $S$  test following by statistically analyzing with Student's  $t$  test or nonparametric test, and enumeration data (HS, aura, and SGS) were assessed using chi-square test, with  $P$ -value  $< .05$  considered statistically significant. For the current research, scikit-learn (Version: 0.20) was used to perform the LASSO, SVMs, RFE, and logistic regression. The predictability of the model was evaluated by the area under the receiver operating characteristic (ROC) curve (AUC). The Youden's index was defined for all points of a ROC curve, the maximum value of which may be used as a criterion for screening the optimum cutoff point. The nomogram was constructed using the rms package in R (Version: 3.5.1). All statistical tests were two-sided.

### 3 | RESULTS

#### 3.1 | Characteristics of individuals

A total of 156 participants (including 10 patients with drug-resistant epilepsy and 20 patients with drug-responsive epilepsy and 30 healthy controls in discovery and training phase, 17 drug-resistant, 13 drug-responsive patients and 18 unclassified patients and 48 healthy controls in validation phase) were recruited to our study. No significant differences of age and gender were found in the training and validation set. The duration of seizures in patients with drug-resistant epilepsy was significantly longer than that in patients with drug-responsive epilepsy ( $P = .01$  in training phases and  $P = .04$  in validation phases). The pathology of hippocampal sclerosis (HS) in patients with drug-resistant epilepsy was significantly more than that in patients with drug-responsive epilepsy ( $P = .002$  in training phases and  $P = .026$  in validation phases). Seizure frequency, the existence of aura, and secondarily generalized seizure were not related to drug response of TLE patients. The detailed clinical characteristics of participants were listed in Table 1.

#### 3.2 | Diagnostic model for TLE

To identify the TLE-associated CpGs, we first studied the global methylation profiles in the DNA of whole peripheral blood obtained from 30 TLE patients and 30 healthy controls. Epigenome-wide association identified 237 DMCs associated with TLE at  $P < .001$  by logistic regression (Figure 1A, Table S1).

LASSO algorithm was used to select the most significant CpGs from 237 candidate CpGs. We then used SVM algorithm to further narrow down the marker numbers. As a result, eight CpGs were identified (cg25838818, cg27564766, cg07782795, cg09383187, cg09293614, cg09270525, cg09197288, and cg08664849), corresponding to *SULT1C2*, *TP73*, *BAIAP2*, *CLIP2*, *MUM1*, *PTPRN2*, *IFI27L1*,

and *TBC1D24*, respectively (Figure 1B, Table S2). Subsequent validation, using a separate validation cohort ( $n = 96$ ), yielded a sensitivity of 71%, a specificity of 73%, and an accuracy of 77%, with an AUC of 0.81 to detect TLE (Figure 1C, Table S3).

#### 3.3 | Prediction model for drug response

T test was used to analyze the DMCs between 10 drug-resistant and 20 drug-responsive patients, which identified 99 DMCs at  $P < .005$ . (Figure 2A, Table S4).

99 DMCs were analyzed by SVM-RFE algorithm to select significant CpGs. Logistic regression was used to further narrow CpGs. Six CpGs were identified (cg15999964, cg08768218, cg11954680, cg17706086, cg21761639, and cg26119877), corresponding to *ZNF608*, *DLC1*, *PCDHA*, *MEST*, and *SLC25A21*, respectively (cg21761639 has no corresponding gene) (Figure 2B; Table S2).

To better investigate the performance of CpGs signature in predicting drug response, a methylation risk score was built with the coefficients weighted by the logistic regression model in the validation cohort (17 drug-resistant, 13 drug-responsive). The methylation risk score was calculated as follows: risk score =  $19.3 * cg15999964 - 43.5 * cg08768218 + 54.9 * cg11954680 + 26.3 * cg17706086 + 66.8 * cg21761639 + 29.9 * cg26119877 - 86.3$ , with a cutoff value of 0.78. Applying the model yielded a sensitivity of 77%, a specificity of 71%, and an accuracy of 73%, with an AUC of 0.79 in the validation cohort, to distinguish drug-responsive from drug-resistant patients (Figure 2C, Table S5).

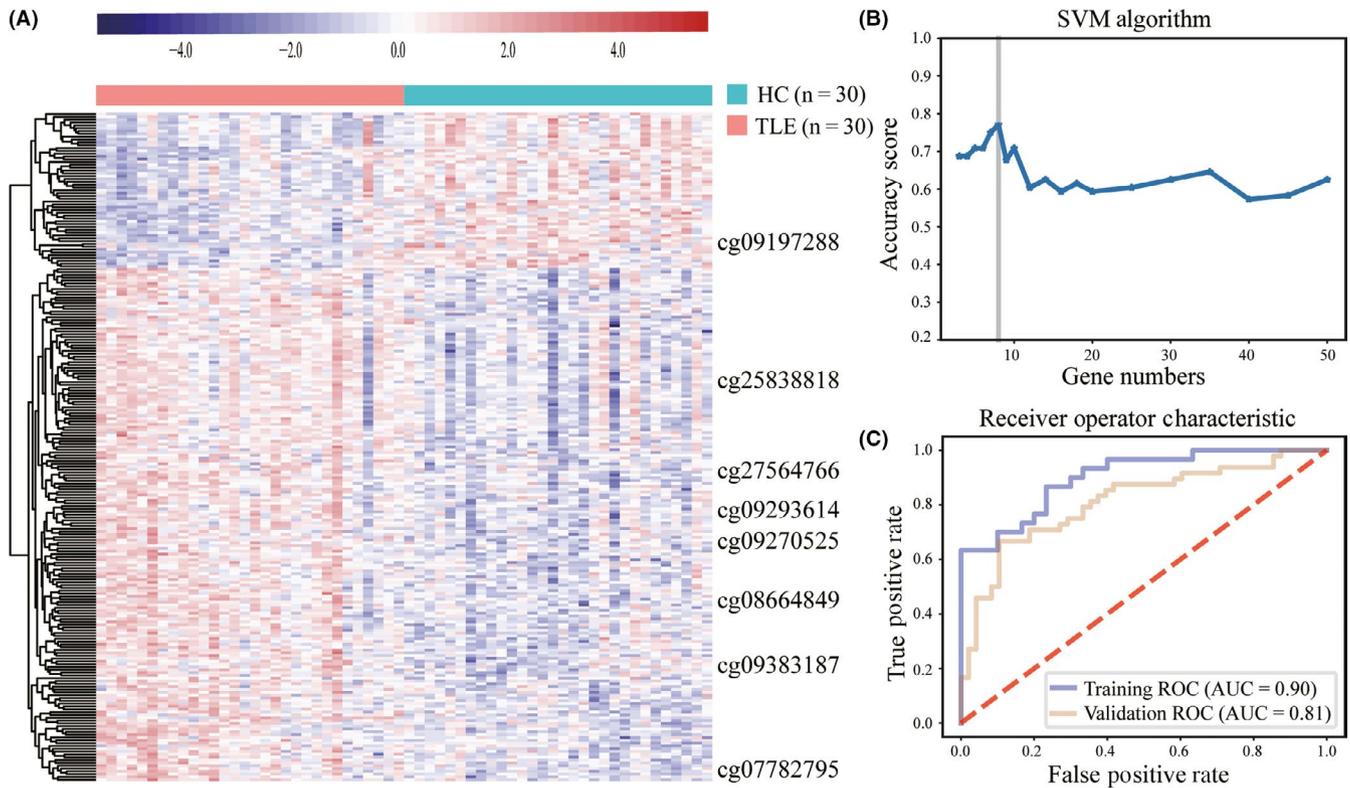
#### 3.4 | Building a predictive nomogram

We performed the multivariate analysis of the methylation risk score and clinicopathological characteristics with drug response in the entire TLE cohort (Figure S1). The methylation risk score and HS were

**TABLE 1** Clinical participants of individuals

	Training set				Validation set			
	Drug-responsive	Drug-resistant	P-value	Control	Drug-responsive	Drug-resistant	P-value	Control
No.	20	10		30	13	17		48
Age, mean $\pm$ SD (y)	28.6 $\pm$ 10.9	31.4 $\pm$ 16.5	.65	31.3 $\pm$ 10.3	18.8 $\pm$ 10.2	31.6 $\pm$ 11.0	.32	31.5 $\pm$ 10.3
Female: male	7:13	5:5	.46	12:18	8:5	9:8	.72	25:23
Disease course, mean(range) (y)	5 (1-24)	13 (4-29)	.01	NA	7 (1-13)	12 (1-26)	.04	NA
Seizure frequency, mean(range) (/ month)	3 (0.1-120)	1 (0.3-90)	.62	NA	2 (0.01-75)	7.5 (1-300)	.28	NA
HS	2 (10%)	7 (70%)	.002	NA	2 (15%)	10 (59%)	.026	NA
Aura	13 (65%)	5 (50%)	.46	NA	7 (54%)	10 (59%)	1.0	NA
SGS	12 (60%)	5 (50%)	.71	NA	10 (77%)	14 (82%)	1.0	NA

Abbreviations: HS, hippocampal sclerosis; NA, not applicable; SD, standard deviation; SGS, secondarily generalized seizure.



**FIGURE 1** Screening and validation of disease-related CpGs. A, Cluster analysis of 237 DMCs associated with TLE at  $P < .001$  by logistic regression in the training cohort. B, SVM algorithms in the training cohort. C, Receiver operator characteristic curve of 8 significant CpGs prediction of TLE patients or healthy controls. The area under the ROC curve in training cohort was 0.90 and 0.81 for validation cohort

significantly associated with drug response. To develop a clinically applicable method that could predict drug-resistance probability of an individual TLE patient, a nomogram was used to build a predictive model in the entire TLE cohort, taking into consideration clinicopathological factors (Figure 3). The predictors included methylation risk score, disease course, seizure frequency, and HS. Applying the model yielded a sensitivity of 94% and a specificity of 89%, with an AUC of 0.96 in the entire TLE patient cohort, to distinguish drug-responsive from drug-resistant patients (Figure 4). The calibration plots for drug-response nomogram model were predicted well in the entire TLE patient cohort. (Figure S2).

### 3.5 | Bioinformatics analysis

All biomarker CpGs corresponding genes were uploaded to IPA for the canonical pathway and diseases functions analysis, and network generation for defined molecular interactions. Results were visualized as networks (Figure S3, Figure S4) and ranked as diseases functions and canonical pathways involved (Table S6). The IPA analysis showed that “cell death and survival” and “cellular development” were the top-ranked diseases functions. *DLC1*, *IFI27L1*, *TP73*, and *PTPRN2* involve in “cell death and survival” and *TBC1D24*, *BAIAP2*, *CLIP2*, and *SULT1C2* involve in “neurological disease.” Furthermore, “DNA methylation and transcriptional repression signaling” were the top-ranked canonical pathways of 6 drug response-related CpGs

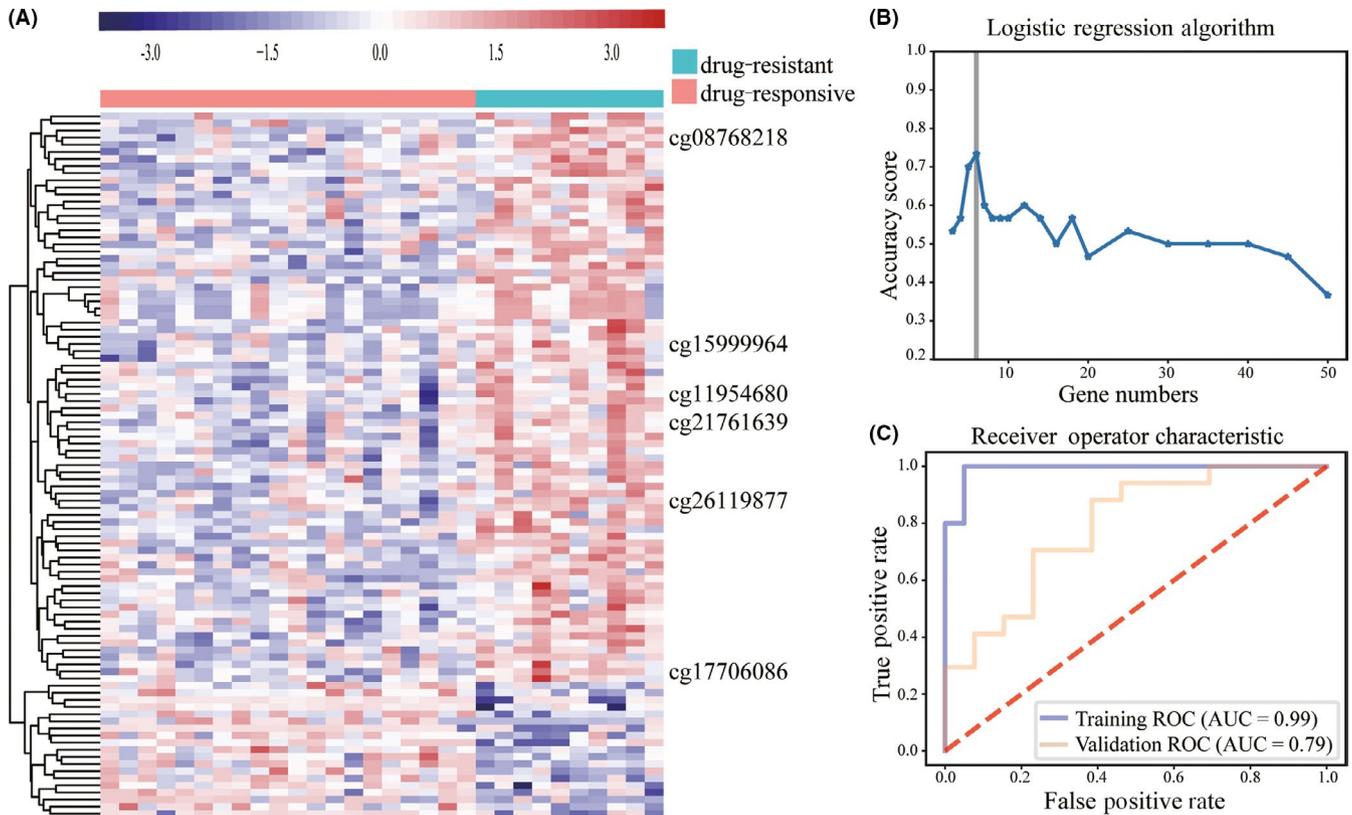
corresponding genes, and *BAIAP2* involves in “axonal guidance signaling” pathway. Notably, *DLC1* is closely related *DNMT* gene (*DNMT1* and *DNMT3B*) in the gene-interaction network of 6 drug response-related CpGs corresponding genes.

### 3.6 | Cross-validation of methylation with bisulfite pyrosequencing

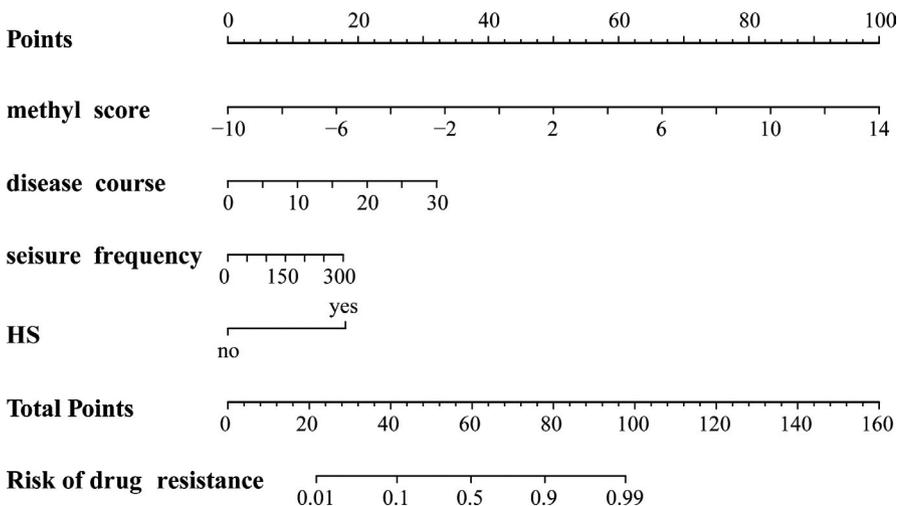
To evaluate the accuracy of DNA methylation data from methylation beadchip, a subset of CpG loci was selected for additional methylation validation by the pyrosequencing. Blood DNA samples of TLE patients ( $n = 10$ ) and controls ( $n = 10$ ) were subjected to methylation detection at 2 loci (*cg25838818*, *cg27564766*), and drug-responsive TLE ( $n = 10$ ) and drug-resistant TLE ( $n = 10$ ) were subjected to methylation detection at 2 loci (*cg11954680* and *cg26119877*). Pyrosequencing revealed methylation of *cg25838818*, *cg27564766*, *cg11954680*, and *cg26119877* was correlated with the data from beadchip array (Figure 5A-D).

## 4 | DISCUSSION

In this study, we used the methylation array to screen differential CpGs and selected significant CpGs by applying machine learning algorithms in the training cohort. Subsequently, we validated the



**FIGURE 2** Screening and validation of drug response–related CpGs. A, Cluster analysis of 99 DMCs associated with TLE drug response at  $P < .005$  by *t* test in the training cohort. B, Logistic regression algorithms in the training cohort. C, Receiver operator characteristic curve of 6 significant CpGs prediction of TLE patients drug-responsive or drug-resistant. The area under the ROC curve in training cohort was 0.99 and 0.79 for validation cohort

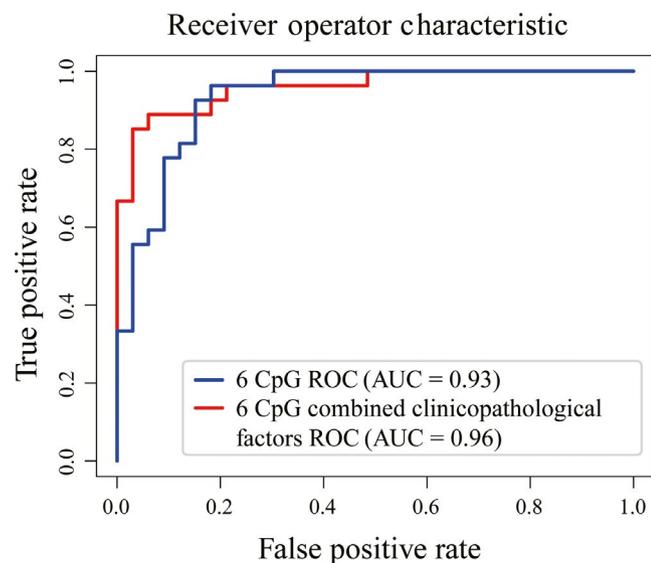


**FIGURE 3** Nomogram to predict the drug response in the entire TLE patient cohort. The nomogram is used by adding up the points identified on the scale for four variables. The sum is located on the “Total points” scale, and a line is drawn downward axes to determine the risk of resistance

candidate CpGs and built the methylation-based signature in the validation cohort. Finally, a nomogram comprising integrated DNA methylation risk score and clinicopathological data was generated to predict the drug response in the entire patient cohort. Our data indicated that the methylation-based signature could define human TLE and predict drug-resistant. The methylation-based biomarker may have clinical applications for individualized diagnosis and treatment outcome prediction for patients with TLE. This study introduced a

methodological framework to screen and validate biomarker and demonstrates the ability to use machine learning as a potential clinical tool for epilepsy diagnosis and drug-response prediction after more comprehensive validation.

While great efforts have been made in understanding the underlying pathogenic and drug-resistant mechanisms of epilepsy, there are no existing treatments to prevent or disease-modify the development.<sup>10</sup> It is believed that the complex and multifactorial features



**FIGURE 4** Receiver operator characteristic curve of 6 significant CpGs combined and not combined clinicopathological factors prediction of drug-responsive or drug-resistant in the entire TLE patient cohort

of epilepsy have led to hampered progress in these areas.<sup>10</sup> Recently, epigenetics as a mediator of gene-environment interactions had appealed to growing interest in understanding the potential role in complex diseases.<sup>42,43</sup> Notably, DNA methylation is attractive to explain the underlying epileptogenesis and pharmacoresistance mechanism in chronic human epilepsy.<sup>27,44-46</sup> Exploring peripheral DNA methylation alterations in epilepsy is considered a direction with translational significance, given that blood samples are available in most clinical settings. Therefore, this study has several strengths should be noted. First, our focus on the specific epilepsy syndrome (TLE), the comparatively homogeneous groups of patients, advantages over the study involving different epilepsy phenotypes. This study features a genome-wide, dual-platform approach to screen TLE biomarkers followed by interrogation of several CpGs in a cohort of samples. Moreover, we replicated these DNA methylation signatures in another independent cohort for validation, and the machine learning model was performed well.

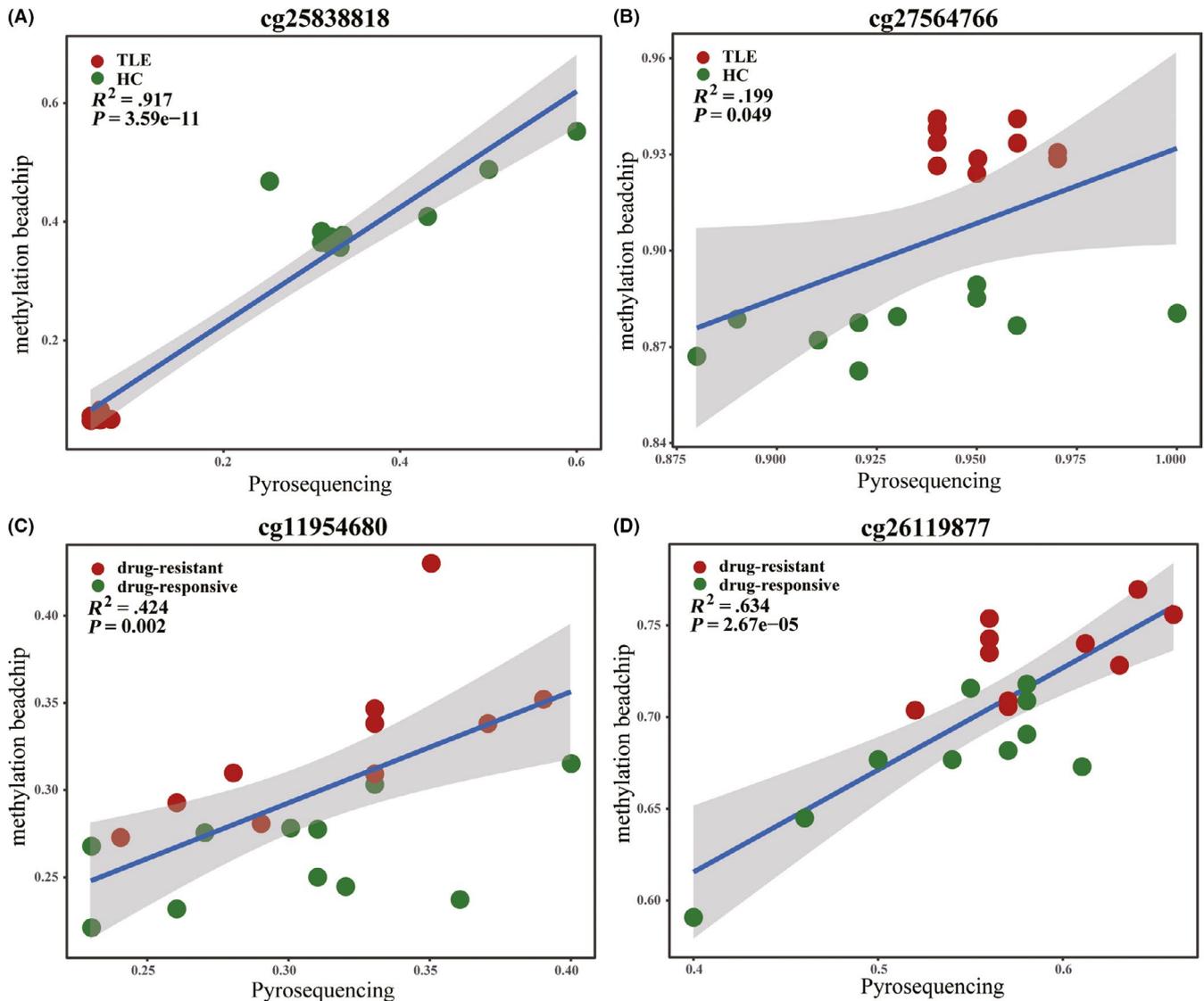
There are additional mechanistic links between all biomarker CpGs corresponding genes and epilepsy. For example, Tre2/Bub2/Cdc16 (TBC)1 domain family member 24 (*TBC1D24*) gene is one of the more recently discovered pathogenic mutations of familial epilepsy, of which associated disorders range from severe epileptic encephalopathy to nonsyndromic hearing loss.<sup>47,48</sup> *TBC1D24* has been implicated in normal neural development and survival and plays an essential role in neurotransmission and presynaptic function.<sup>49</sup> In sporadic mesial TLE-HS, whole-exome sequencing has identified nonsynonymous de novo variants in *BAIAP2* gene, of which knockout model might produce aberrant neurological phenotypes listed by Mouse Genome Informatics (MGI).<sup>50</sup> Interestingly, adenosine treatment reversed the DNA hypermethylation (including *Sult1c2* gene) status in the brain of TLE model, inhibited sprouting of mossy fibers in the hippocampus, and prevented epileptogenesis.<sup>51</sup>

Further characterization of molecules such as *TBC1D24*, *BAIAP2*, and *SULT1C2* will provide new insights into TLE development and progression.

In addition, we noted that drug response-related CpGs corresponding genes closely related to DNA methylation, which implicated that DNA methylation play an essential role in pharmacoresistance mechanisms of epilepsy. Tumor suppressor gene deleted in liver cancer 1 (*DLC1*) is shown to induce apoptosis, frequently silenced by methylation and negative correlation with DNMT expression.<sup>52,53</sup> Interestingly, previous research found increased expression of DNMT1 and DNMT3A in patients with intractable TLE.<sup>22</sup> Hypermethylation of gene promoters was also the predominant effect in TLE patients and rodent models as well.<sup>24,27-29</sup> Given the well studied of epigenetic pathomechanisms underlying drug resistance in cancer, Kobow proposed that the methylation hypothesis of pharmacoresistance could open such new avenue in the field of epilepsy.<sup>45</sup>

We produced a nomogram including DNA methylation risk score, disease course, seizure frequency, and HS for estimation of individualized outcomes of the drug response in TLE patients. The AUC of predictive model is as high as 0.96, which suggests that this model is promising to be applicable in clinical practice. In the previous research, some seizure-related characteristics, including the prior number of seizures and disease course, were reported to be related with the risk of drug resistance.<sup>4,9</sup> Another factor strongly linked to the increased risk of intractable epilepsy is HS, which is consistent with the findings in our study. However, limited studies applied combinatorial biomarker signatures to predict drug response in epilepsy patients. A model composed of clinical variables (the presence of HS) in combination with genetic information (SNP genotypes located in 11 genes influencing drug transport and metabolism) improved predictive accuracy for medical intractability in mesial TLE.<sup>54</sup> With the advent of high-throughput technologies and the availability of multidimensional data sets, we suggest the need to combine compound molecular approaches to achieve higher predictive performance for clinical usefulness and better comprehend the knowledge about the relevant underlying pathomechanism.

The study also has certain limits and constraints that should be noted when interpreting the results. First of all, although our classification results were promising, we should point out that high dimensional data with small sample size may result in misclassifications and biased predictors. A larger set of patients can enhance the robustness of the predictive model. Second, due to the single-center study, the results of this pilot study warrant further validation in samples from several neurological centers. Third, TLE patients participated in this study were all being treated with antiepileptic medication, and it is still unknown whether these might affect the DNA methylation status. Furthermore, other confounding factors, such as cellular composition in whole blood, should also be taken into consideration.<sup>55</sup> Fourth, two different platforms of methylation data set were hired that is 450K array for training while 850K array for validation. Several studies have reported that overall correlations of matched samples running both on the 450K and 850K array were quite high



**FIGURE 5** Cross-validation of DNA methylation with the pyrosequencing. Shown are degrees of methylation of 4 CpG loci reported by methylation BeadChip (Y axis, ratio) and pyrosequencing (X axis, ratio) assays. For cg25838818 (A), cg27564766 (B), cg11954680 (C), and cg26119877 (D), the degrees of methylation detected by the two methods were positively correlated ( $P < .05$ ) in reference to individual samples

( $r > .90$  for all assessed samples).<sup>56,57</sup> Last, longitudinal study including the patients before antiepileptic medication is recommended to warrant clinical significance of the predictive biomarkers. The biologic mechanisms of the candidate markers are still little known, and thorough in vivo and vitro experiments are also needed to future investigate.

## 5 | CONCLUSIONS

For the first time, we demonstrated that DNA methylation signature could define human TLE and compound with clinicopathological factors to improve the prediction of response to drug treatment. Furthermore, this study introduced a methodological framework to screen and validate biomarker and demonstrated the ability to use machine learning as a potential clinical investigative tool. Despite the

limited pathomechanism contributions, we highlight the utilization of promising biomarkers in clinical practice for decision-making.

## ACKNOWLEDGMENTS

Foremost, we thank the patients, their families, and healthy controls for their cooperation and participation in this study. This study was supported financially by the National Natural Science Foundation of China (No.81671299 and No.81974206 to BX, No.81671300 to LLL and No.81701182 to YJF), Project from Hunan Provincial Science and Technology Department (No.2019SK1012 to BX), Clinical Research Foundation of Xiangya Hospital (No.2016L08 to LLL), and Fundamental Research Funds for the Central Universities of Central South University (No. 2018zzts248 to WBX).

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ETHICAL APPROVAL

The study was approved by the Ethics Committee of Central South University, Xiangya School of Medicine and the affiliated Xiangya Hospital (201303120). Informed consent was obtained from the parents/legal guardians of all patients.

## ORCID

Wenbiao Xiao  <https://orcid.org/0000-0002-8922-4654>

Bo Xiao  <https://orcid.org/0000-0001-5204-1902>

Hongyu Long  <https://orcid.org/0000-0001-5641-8850>

## REFERENCES

- Tellez-Zenteno JF, Hernandez-Ronquillo L. A review of the epidemiology of temporal lobe epilepsy. *Epilepsy Res Treat*. 2012;2012:630853.
- Blumcke I, Thom M, Aronica E, et al. International consensus classification of hippocampal sclerosis in temporal lobe epilepsy: a task force report from the ILAE commission on diagnostic methods. *Epilepsia*. 2013;54(7):1315-1329.
- Janszky J, Janszky I, Schulz R, et al. Temporal lobe epilepsy with hippocampal sclerosis: predictors for long-term surgical outcome. *Brain*. 2005;128(Pt 2):395-404.
- Bjorke AB, Nome CG, Falk RS, Gjerstad L, Tauboll E, Heuser K. Evaluation of long-term antiepileptic drug use in patients with temporal lobe epilepsy: assessment of risk factors for drug resistance and polypharmacy. *Seizure*. 2018;61:63-70.
- Kanner AM. Management of psychiatric and neurological comorbidities in epilepsy. *Nat Rev Neurol*. 2016;12(2):106-116.
- Raof R, Jimenez-Mateos EM, Bauer S, et al. Cerebrospinal fluid microRNAs are potential biomarkers of temporal lobe epilepsy and status epilepticus. *Sci Rep*. 2017;7(1):3328.
- Wang J, Yu JT, Tan L, et al. Genome-wide circulating microRNA expression profiling indicates biomarkers for epilepsy. *Sci Rep*. 2015;5:9522.
- Kwan P, Brodie MJ. Early identification of refractory epilepsy. *N Engl J Med*. 2000;342(5):314-319.
- Huang L, Li S, He D, Bao W, Li L. A predictive risk model for medical intractability in epilepsy. *Epilepsy Behav*. 2014;37:282-286.
- Pitkanen A, Loscher W, Vezzani A, et al. Advances in the development of biomarkers for epilepsy. *Lancet Neurol*. 2016;15(8):843-856.
- Pitkanen A, Ekolle NX, Lapinlampi N, Puhakka N. Epilepsy biomarkers - Toward etiology and pathology specificity. *Neurobiol Dis*. 2019;123:42-58.
- Engel JJ, Pitkanen A, Loeb JA, et al. Epilepsy biomarkers. *Epilepsia*. 2013;54(Suppl 4):61-69.
- Raof R, Bauer S, El NH, et al. Dual-center, dual-platform microRNA profiling identifies potential plasma biomarkers of adult temporal lobe epilepsy. *Ebiomedicine*. 2018;38:127-141.
- Wang J, Tan L, Tan L, et al. Circulating microRNAs are promising novel biomarkers for drug-resistant epilepsy. *Sci Rep*. 2015;5:10201.
- Koch A, Joosten SC, Feng Z, et al. Analysis of DNA methylation in cancer: location revisited. *Nat Rev Clin Oncol*. 2018;15(7):459-466.
- Kinde B, Gabel HW, Gilbert CS, Griffith EC, Greenberg ME. Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. *Proc Natl Acad Sci USA*. 2015;112(22):6800-6806.
- Bestor TH. The DNA methyltransferases of mammals. *Hum Mol Genet*. 2000;9(16):2395-2402.
- Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell*. 2007;128(4):669-681.
- Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13(7):484-492.
- Belhedi N, Perroud N, Karege F, Vessaz M, Malafosse A, Salzmann A. Increased CPA6 promoter methylation in focal epilepsy and in febrile seizures. *Epilepsy Res*. 2014;108(1):144-148.
- Unal Y, Kara M, Genc F, et al. The methylation status of NKCC1 and KCC2 in the patients with refractory temporal lobe epilepsy. *Idogyogy Sz*. 2019;72(5-6):181-186.
- Zhu Q, Wang L, Zhang Y, et al. Increased expression of DNA methyltransferase 1 and 3a in human temporal lobe epilepsy. *J Mol Neurosci*. 2012;46(2):420-426.
- Kobow K, Jeske I, Hildebrandt M, et al. Increased reelin promoter methylation is associated with granule cell dispersion in human temporal lobe epilepsy. *J Neuropathol Exp Neurol*. 2009;68(4):356-364.
- Kobow K, Kaspi A, Harikrishnan KN, et al. Deep sequencing reveals increased DNA methylation in chronic rat epilepsy. *Acta Neuropathol*. 2013;126(5):741-756.
- Debski KJ, Pitkanen A, Puhakka N, et al. Etiology matters - genomic DNA methylation patterns in three rat models of acquired epilepsy. *Sci Rep*. 2016;6:25668.
- Miller-Delaney SF, Das S, Sano T, et al. Differential DNA methylation patterns define status epilepticus and epileptic tolerance. *J Neurosci*. 2012;32(5):1577-1588.
- Miller-Delaney SF, Bryan K, Das S, et al. Differential DNA methylation profiles of coding and non-coding genes define hippocampal sclerosis in human temporal lobe epilepsy. *Brain*. 2015;138(Pt 3):616-631.
- Long HY, Feng L, Kang J, et al. Blood DNA methylation pattern is altered in mesial temporal lobe epilepsy. *Sci Rep*. 2017;7:43810.
- Xiao W, Cao Y, Long H, et al. Genome-wide DNA methylation patterns analysis of noncoding RNAs in temporal lobe epilepsy patients. *Mol Neurobiol*. 2018;55(1):793-803.
- Urduingio RG, Sanchez-Mut JV, Esteller M. Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. *Lancet Neurol*. 2009;8(11):1056-1072.
- Li L, Choi JY, Lee KM, et al. DNA methylation in peripheral blood: a potential biomarker for cancer molecular epidemiology. *J Epidemiol*. 2012;22(5):384-394.
- Sina AA, Carrascosa LG, Trau M. DNA methylation-based point-of-care cancer detection: challenges and possibilities. *Trends Mol Med*. 2019;25(11):955-966.
- Wick W, Weller M, van den Bent M, et al. MGMT testing—the challenges for biomarker-based glioma treatment. *Nat Rev Neurol*. 2014;10(7):372-385.
- Bell EH, Zhang P, Fisher BJ, et al. Association of MGMT promoter methylation status with survival outcomes in patients with high-risk glioma treated with radiotherapy and temozolomide: an analysis from the NRG oncology/RTOG 0424 trial. *Jama Oncol*. 2018;4(10):1405-1409.
- Church TR, Wandell M, Lofton-Day C, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut*. 2014;63(2):317-325.
- Qiu J, Peng B, Tang Y, et al. CpG methylation signature predicts recurrence in early-stage hepatocellular carcinoma: results from a multicenter study. *J Clin Oncol*. 2017;35(7):734-742.
- Widschwendter M, Zikan M, Wahl B, et al. The potential of circulating tumor DNA methylation analysis for the early detection and management of ovarian cancer. *Genome Med*. 2017;9(1):116.
- Li Y, Gu J, Xu F, Zhu Q, Ge D, Lu C. Novel methylation-driven genes identified as prognostic indicators for lung squamous cell carcinoma. *Am J Transl Res*. 2019;11(4):1997-2012.
- Wang L, Shi J, Huang Y, et al. A six-gene prognostic model predicts overall survival in bladder cancer patients. *Cancer Cell Int*. 2019;19:229.
- Maksimovic J, Gordon L, Oshlack A. SWAN: subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol*. 2012;13(6):R44.

41. Delaney C, Garg SK, Yung R. Analysis of DNA methylation by pyrosequencing. *Methods Mol Biol.* 2015;1343:249-264.
42. Gal-Yam EN, Saito Y, Egger G, Jones PA. Cancer epigenetics: modifications, screening, and therapy. *Annu Rev Med.* 2008;59:267-280.
43. Qureshi IA, Mehler MF. Understanding neurological disease mechanisms in the era of epigenetics. *Jama Neurol.* 2013;70(6):703-710.
44. Kobow K, Blumcke I. The methylation hypothesis: do epigenetic chromatin modifications play a role in epileptogenesis? *Epilepsia.* 2011;52(Suppl 4):15-19.
45. Kobow K, El-Osta A, Blumcke I. The methylation hypothesis of pharmacoresistance in epilepsy. *Epilepsia.* 2013;54(Suppl 2):41-47.
46. Kobow K, Blumcke I. Epigenetics in epilepsy. *Neurosci Lett.* 2018;667:40-46.
47. Balestrini S, Milh M, Castiglioni C, et al. TBC1D24 genotype-phenotype correlation: epilepsies and other neurologic features. *Neurology.* 2016;87(1):77-85.
48. Lozano R, Herman K, Rothfuss M, et al. Clinical intrafamilial variability in lethal familial neonatal seizure disorder caused by TBC1D24 mutations. *Am J Med Genet A.* 2016;170(12):3207-3214.
49. Finelli MJ, Aprile D, Castroflorio E, et al. The epilepsy-associated protein TBC1D24 is required for normal development, survival and vesicle trafficking in mammalian neurons. *Hum Mol Genet.* 2019;28(4):584-597.
50. Wong J, Gui H, Kwok M, et al. Rare variants and de novo variants in mesial temporal lobe epilepsy with hippocampal sclerosis. *Neurol Genet.* 2018;4(3):e245.
51. Williams-Karnesky RL, Sandau US, Lusardi TA, et al. Epigenetic changes induced by adenosine augmentation therapy prevent epileptogenesis. *J Clin Invest.* 2013;123(8):3552-3563.
52. Seng TJ, Low JS, Li H, et al. The major 8p22 tumor suppressor DLC1 is frequently silenced by methylation in both endemic and sporadic nasopharyngeal, esophageal, and cervical carcinomas, and inhibits tumor cell colony formation. *Oncogene.* 2007;26(6):934-944.
53. Fang QL, Yin YR, Xie CR, et al. Mechanistic and biological significance of DNA methyltransferase 1 upregulated by growth factors in human hepatocellular carcinoma. *Int J Oncol.* 2015;46(2):782-790.
54. Silva-Alves MS, Secolin R, Carvalho BS, et al. A prediction algorithm for drug response in patients with mesial temporal lobe epilepsy based on clinical and genetic information. *PLoS ONE.* 2017;12(1):e169214.
55. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol.* 2014;15(2):R31.
56. Logue MW, Smith AK, Wolf EJ, et al. The correlation of methylation levels measured using Illumina 450K and EPIC BeadChips in blood samples. *Epigenomics-UK.* 2017;9(11):1363-1371.
57. 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.

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Xiao W, Liu C, Zhong K, et al. CpG methylation signature defines human temporal lobe epilepsy and predicts drug-resistant. *CNS Neurosci Ther.* 2020;26:1021-1030. <https://doi.org/10.1111/cns.13394>