ARTICLE

Genomic DNA Methylation Signatures Enable Concurrent Diagnosis and Clinical Genetic Variant Classification in Neurodevelopmental Syndromes

Erfan Aref-Eshghi,[1,2] David I. Rodenhiser,[3] Laila C. Schenkel,[1] Hanxin Lin,[1,2] Cindy Skinner,[4] Peter Ainsworth,[1,2] Guillaume Paré,[5] Rebecca L. Hood,[6] Dennis E. Bulman,[7] Kristin D. Kernohan,[7] Care4Rare Canada Consortium, Kym M. Boycott,[7] Philippe M. Campeau,[8] Charles Schwartz,[4] and Bekim Sadikovic[1,2,*]

Pediatric developmental syndromes present with systemic, complex, and often overlapping clinical features that are not infrequently a consequence of Mendelian inheritance of mutations in genes involved in DNA methylation, establishment of histone modifications, and chromatin remodeling (the "epigenetic machinery"). The mechanistic cross-talk between histone modification and DNA methylation suggests that these syndromes might be expected to display specific DNA methylation signatures that are a reflection of those primary errors associated with chromatin dysregulation. Given the interrelated functions of these chromatin regulatory proteins, we sought to identify DNA methylation epi-signatures that could provide syndrome-specific biomarkers to complement standard clinical diagnostics. In the present study, we examined peripheral blood samples from a large cohort of individuals encompassing 14 Mendelian disorders displaying mutations in the genes encoding proteins of the epigenetic machinery. We demonstrated that specific but partially overlapping DNA methylation signatures are associated with many of these conditions. The degree of overlap among these epi-signatures is minimal, further suggesting that, consistent with the initial event, the downstream changes are unique to every syndrome. In addition, by combining these epi-signatures, we have demonstrated that a machine learning tool can be built to concurrently screen for multiple syndromes with high sensitivity and specificity, and we highlight the utility of this tool in solving ambiguous case subjects presenting with variants of unknown significance, along with its ability to generate accurate predictions for subjects presenting with the overlapping clinical and molecular features associated with the disruption of the epigenetic machinery.

Introduction

Genes encoding the epigenetic protein machinery that read, write, and erase post-translational signals on DNA and histones and remodel chromatin are implicated in a wide range of constitutional neurodevelopmental disorders.[1–3] The pathogenesis of such disorders is likely caused by the downstream events orchestrated by the primary functional defect in these proteins of the so-called epigenetic machinery.[1–3] Furthermore, specific mutations in these readers, writers, erasers, and chromatin remodelers are linked to the variability in clinical phenotype seen in their associated disorders. It is well established that histone modifications overlap and interact with genomic DNA methylation to affect chromatin remodeling,[4] and thus, that mutations in the genes that are involved in histone modifications are expected to have an impact within the DNA methylome. Supporting this concept, we have previously reported DNA methylation epigenetic (epi-) signatures in the peripheral blood of subjects carrying mutations in genes involved in chromatin regulation, including the mutations in *SRCAP* (MIM: 611421) causing Floating-Harbor syndrome (MIM: 136140),[5] *DNMT1* (MIM: 126375) resulting in adult-onset autosomal-dominant cerebellar ataxia, deafness, and narcolepsy (ADCA-DN [MIM: 604121]),[6] and *ATRX* (MIM: 300032), which is responsible for the alpha thalassemia/mental retardation X-linked (ATRX) syndrome (MIM: 300448).[7] Epi-signatures in subjects with Sotos (MIM: 117550), CHARGE (MIM: 214800), and Kabuki (MIM: 147920) syndromes have also been reported.[8–10]

While the number of developmental and cancer-related conditions for which a DNA methylation epi-signature has been reported is increasing, the extent of overlap or distinction of these epi-signatures is not clear. Clinical overlap is a common finding in the diseases that result from such defects, and it is postulated that mechanistic overlap could be a basis for such phenotypic similarity. This is further acknowledged by noting that all of the proteins associated with these conditions regulate the epigenome through complex interactions with each other. Hence, the question has been raised as to how accurately one can use these epi-signatures as a tool in the molecular diagnosis of these conditions.[9,10] This is particularly

[1]Department of Pathology and Laboratory Medicine, Western University, London, ON N6A5C1, Canada; [2]Molecular Genetics Laboratory, Molecular Diagnostics Division, London Health Sciences Centre, London, ON N6A5W9, Canada; [3]Departments of Pediatrics, Biochemistry and Oncology, Western University and Children's Health Research Institute, London, ON N6A5C1, Canada; [4]Greenwood Genetics Center, Greenwood, SC 29646, USA; [5]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON L8S4L8, Canada; [6]Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, ON K1H8M5, Canada; [7]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON K1H5B2, Canada; [8]Department of Pediatrics, University of Montreal, Montreal, QC H3T1J4, Canada
*Correspondence: bekim.sadikovic@lhsc.on.ca
https://doi.org/10.1016/j.ajhg.2017.12.008.

Table 1. Structure and Demographics of the Study Cohort

			Discovery-Training Cohort			Control Cohort					
Syndrome	Total	Testing Cohort	No. of Individuals	Percentage Female	Mean Age ± SD	No. of Individuals	Percentage Female	Mean Age ± SD	Probes Passing QC	Probes Found	DMRs Found
Rett	17	4	13	92%	7.6 ± 11	52	92%	7.4 ± 10	448,775	N/A	no
Saethre-Chotzen	25	6	19	63%	10.6 ± 11	76	63%	10.11 ± 10	442,079	N/A	no
Weaver	7	2	5	40%	age unknown	20	40%	–	455,427	N/A	no
Coffin Siris	9	3	6	50%	5.9 ± 6	24	50%	5.8 ± 5	453,321	N/A	no
Coffin Lowry	11	3	8	12%	9.8 ± 6	32	12%	9.7 ± 6	453,285	N/A	no
ATRX	19	4	15	0%	11.4 ± 7	60	0%	11.3 ± 7	450,748	1,112	41
Floating-Harbor	17	4	13	76%	12.7	52	76%	12.2 ± 11	453,189	1,078	54
Sotos	38	10	28	57%	8.8 ± 4	112	57%	8.9 ± 4	448,131	6,858	1,372
ADCA-DN	5	0	5	40%	age unknown	20	40%	–	330,788	3,562	52
Claes-Jensen	10	2	8	0%	22.5 ± 14.2	32	0%	21.2 ± 13	454,978	698	14
Kabuki	44	11	33	57%	9.5 ± 6	132	57%	9.7 ± 7	401,051	919	31
CHARGE	79	40	39	41%	5.8 ± 6	156	48%	5.9 ± 6	448,876	1,320	18
GTPTS	3	0	3	33%	2.1 ± 5.8	20	33%	2.2 ± 5.1	454,489	707	6
SBBYSS	1	0	1	100%	6.3 ± 0	20	100%	6.5 ± 0.8	456,134	864	6

important in the context of utilizing these epi-signatures as functional evidence to classify variants of uncertain clinical significance.

In the present study, we try to address these questions by concurrently examining 14 Mendelian conditions that result from direct or indirect disruptions of the proteins involved in the regulation of the epigenome. These conditions (see Table 1) include Rett syndrome (MIM: 312750) (methyl-CpG-binding protein 2; *MeCP2* [MIM: 300005]), ADCA-DN (DNA methyltransferase 1; *DNMT1*), Kabuki syndrome (lysine-specific methyltransferase 2D; *KMT2D* [MIM: 147920]), ATRX syndrome (*ATRX*), Sotos syndrome (nuclear receptor binding SET domain protein 1; *NSD1* [MIM: 117550]), Floating-Harbor syndrome (Snf2 related CREBBP activator protein; *SRCAP*), Weaver syndrome (MIM: 277590) (enhancer of zeste 2 polycomb repressive complex 2 subunit; *EZH2* [MIM: 601573]), CHARGE syndrome (chromodomain helicase DNA binding protein 7; *CHD7* [MIM: 608892]), Claes-Jensen syndrome (MIM: 300534) (lysine-specific demethylase 5C; *KDM5C* [MIM: 314690]), Genitopatellar syndrome (GTPTS [MIM: 606170]), and Say-Barber-Biesecker-Young-Simpson syndrome (SBBYSS [MIM: 603736]), both caused by mutations in lysine acetyltransferase 6B (*KAT6B* [MIM: 605880]), and Coffin-Siris syndrome (MIM: 135900) (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily B1 *SMARCB1* [MIM: 601607], and AT-rich interaction domain 1B; *AR1D1B* [MIM: 614556]). These genes are directly involved in epigenomic regulation of the chromatin. We have also included two other conditions that result from mutations in genes that interact with the components of the epigenomic machinery. Saethre-Chotzen syndrome (MIM: 101400) is caused by mutations in the Twist family BHLH transcription factor 1 gene (*TWIST* [MIM: 601622]), encoding a transcription factor that binds to the p300 and p300/CREBBP-associated factor domains of histone acetyltransferases and regulates their activity,[11] and Coffin-Lowry syndrome (MIM: 303600), which results from mutations in ribosomal protein S6 kinase (*RSK2* [MIM: 300075]), a protein required for phosphorylation of histone H3 that regulates chromatin remodeling and gene expression.[12]

In this study, we have identified DNA methylation profiles in peripheral blood samples from a large cohort of individuals who carry mutations in the associated candidate genes responsible for their respective developmental syndromes, including previously unreported signatures in cohorts for GTPTS and SBBYSS (*KAT6B*) and Claes-Jensen syndrome (*KDM5C*). We have also examined the degree of positional overlap between these epi-signatures and have evaluated whether a single classification tool, built on the combined epi-signatures of these conditions, can generate accurate predictions for subjects presenting with the overlapping clinical and molecular features associated with mutations in these genes that play an essential role in the epigenetic machinery.

Material and Methods

Source of Data

This study utilized data and specimens from multiple sources. Peripheral blood DNA samples from subjects with ADCA-DN, Coffin-Siris syndrome, SBBYSS, GTPTS, and Floating-Harbor syndrome were collected from the Care4Rare Canada Consortium. Samples from subjects with clinical characteristics of Kabuki syndrome, ATRX, Saethre-Chotzen syndrome, Coffin-Lowry syndrome, Rett syndrome, Claes-Jensen syndrome, and CHARGE syndrome were collected from the Greenwood Genetic Center (Greenwood, SC, USA). The Kabuki and CHARGE cohorts were supplemented by methylation array files publicly available from GEO (GSE97362).[9] Epigenomic data from subjects with Sotos and Weaver syndromes were downloaded from GEO (GSE74432).[8] All of these subjects had clinical features of the aforementioned syndromes and were screened for mutations in the related genes. The mutation report from every subject was reviewed according to the American College of Medical Genetics Guidelines for interpretation of genomic sequence variants,[13] and only individuals confirmed to carry a pathogenic or likely pathogenic mutation were used to identify epi-signatures (subject-level data are summarized in Table S1 where available). Control subjects were selected from our lab reference cohort, which is composed of individuals with no known aberrant epigenomic change. This reference cohort was previously preselected from a larger cohort of approximately 1,000 individuals across a broad range of age, sex, and ethnicity distribution.

Methylation Array and Quality Assessment

Genomic DNA was extracted from peripheral blood using standard techniques. Following bisulfite conversion, DNA methylation analysis of the samples was performed using the Illumina Infinium bead chip array, according to the manufacturer's protocol at the Genetic and Molecular Epidemiology Laboratory at McMaster University and the London Health Sciences Molecular Genetic Laboratory. Except for a cohort of subjects with CHARGE syndrome ($n = 39$), which was assayed using Illumina Infinium methylation EPIC array, all of the samples were assayed using the HumanMethylation450 bead chip. The two arrays harbor 96% overlap in CpG probes. Methylated and unmethylated intensity data were generated as idat files and imported into R 3.4.0 for analysis. Normalization was performed using Illumina normalization method with background correction using the minfi package. Probes with detection p value > 0.01 were excluded from the downstream analyses. For further quality improvement, probes located on chromosomes X and Y and probes known to contain SNPs at the CpG interrogation, or the single-nucleotide extension, were removed. As an additional quality-control step, the sex of the samples was predicted using the signal intensity of the X and Y chromosomes using the minfi package, and the files representing a discordance between the predicted and labeled sex were not used for identification of the DNA methylation profile. All of the samples were examined for genome-wide methylation density, and those deviating from a bimodal signal distribution were excluded. Files generated by the EPIC array were cast as 450k array, and the same analytical procedures used on the 450k array were applied.

Selection of Discovery/Training and Testing Cohorts and Controls

The identification of disease-specific epi-signatures was performed using a randomly selected 75% subset of the database (discovery/training set) using caTools package. The remaining samples were only used as a testing cohort to assess the performance of the classification model developed later in the study. This procedure was not performed when ≤ 5 samples were available for a disease group (ADCA-DN, GTPTS, and SBBYSS). Given probe differences and technology variations between the two array types (450k and EPIC), it was ensured that the entire discovery/training cohort is assayed using one array (450k). Therefore, all of the CHARGE-affected subjects who were assayed using the 450k array were selected as the discovery/training cohort and the rest were included in the testing cohort. For every disease group in the discovery cohort, a sex- and age-matched control group with a sample size at least four times larger (minimum $n = 20$) was selected from the reference control group using MatchIt package. The methylation profile of each disease group in the discovery cohort was compared with its matched control separately to identify the disease-specific epi-signature (Table 1). Figure 1 represents the flowchart of the study.

Identification of Disease-Specific Methylation Epi-signatures

Analysis was performed using a modification of our previously published protocol.[5–7,10,14] The methylation level for each probe was measured as a beta value, calculated from the ratio of the methylated signals versus the total sum of unmethylated and methylated signals, ranging between 0 (no methylation) and 1 (full methylation). This value was used for biological interpretation and visualization. For statistical analysis, wherever a normal distribution was required (linear regression modeling), beta values were logit transformed to M-values using the following equation: log2(beta/(1-beta)). A linear regression modeling using the limma package was used to identify the differentially methylated probes. The analysis was adjusted for blood cell type compositions predicted using minfi package. The generated p values were moderated using the eBayes function in the limma package and were corrected for multiple testing using Benjamini and Hochberg method. Probes with a corrected p value < 0.01 and a methylation difference greater than 10%–20% were considered significant. The effect size cutoffs (10%–20%) were determined separately for every condition following the examination of the volcano plots generated in every comparison as previously conducted by Butcher et al.[9] The identified probes were examined using an unsupervised hierarchical clustering to ensure their ability in separating the subjects from controls. In the case of the Coffin-Siris syndrome, which was caused by two genes in our dataset, the cohort was first split based on the causing genes and then re-analyzed regardless of the gene.

Control for Batch Effect and Robustness of the Identified Epi-signatures

Given that the study cohort was composed of data generated by multiple centers, different dates, and via various instruments, several measures were utilized to minimize possible batch effects and other sources of variability. These steps involved: (1) whenever the sample size allowed, control subjects were selected from the same batch (as with ADCA-DN); (2) if case subjects and their control subjects in a comparison were assayed in multiple batches (as with Kabuki syndrome samples), the batch variable was included as a confounding factor in the regression model; (3) where a particular sample was assayed using both 450k and EPIC arrays (CHARGE syndrome), only files from 450k were used for identification of the signature; (4) for all other disease cohorts, a

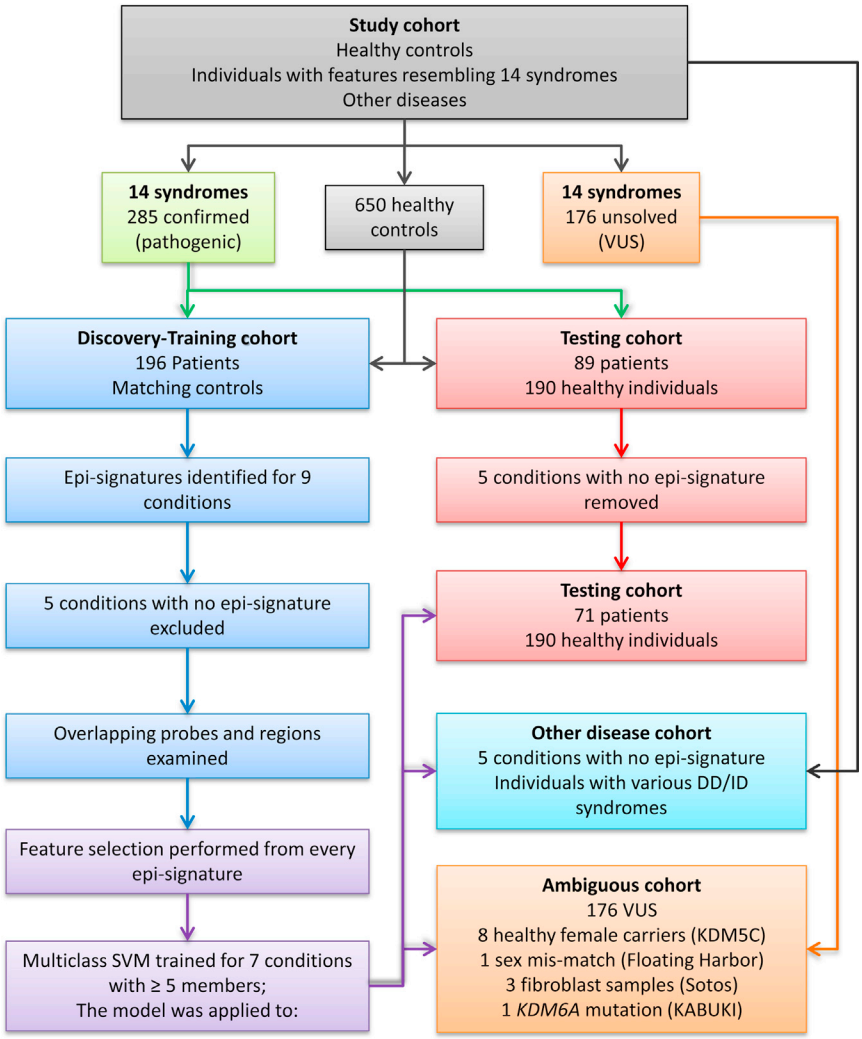

Figure 1. Flowchart of the Study

machine learning model (Support vector machine) was trained on the identified probes, and we examined how well it could distinguish the batch structure from the disease status. Failure to pass this step resulted in the findings being regarded as non-reliable.

Identification of Genomic Regions with Methylation Changes

To identify genomic regions harboring methylation changes (differentially methylated regions [DMRs]), a bump hunting approach was used by the bumphunter package.[15] The analysis considered regions with greater than 10% change in the overall methylation between case and control subjects with gaps no more than 500 bp among neighboring CpGs. As suggested in the package, 1,000 bootstrapping procedure was performed to compute family-wise error rate (FWER). We selected regions containing a minimum of three consecutive probes and FWER < 0.01. The identified regions were mapped to CpG islands and coding genes. Gviz package was used for visualization of the DMRs.

Assessment of the Overlap between the Epi-signatures

Probes and regions differentially methylated in each disease group were examined to identify potential overlap. The number of probes shared in more than one disease group was visualized using a circos plot. Probes that were shared in >2 and 3 disease groups were used to measure the pairwise correlations in the disease cohort. Calculated correlation coefficients were visualized using a correlation plot. Genomic regions harboring differentially methylated probes were assessed for overlap using GenomicRanges package, and regions found in more than one disease type were reported. Functional annotation clustering and gene set enrichment analysis was performed using missMethyl and ReactomePA packages for the genes harboring the shared probes.

Construction and Validation of a Multi-class Prediction Model

The identified signatures were used to build a classification model with the ability to concurrently assess a given methylation profile belonging to any of the disease groups in the study. Caret package was used for feature selection from every signature. First, a receiver operating characteristic curve analysis was performed to identify the most differentiating probes. Those probes with an area under the curve above 0.8 were retained. Next, pairwise correlations among the remaining probes were measured to identify and exclude the redundant signals with R-squared > 0.8. A multi-class support vector machine (SVM) with linear kernel was trained on the remaining probes using e1071 package. To determine the best hyperparameters and to measure the accuracy of the model, a 10-fold cross-validation was performed. In this process, the training set was divided into ten folds. Nine folds were used for training the model and one fold for testing. After repeating this iteration for all of the ten folds, the mean accuracy was calculated and the hyperparameters with the optimal performance were selected. For every sample, the model was set to generate multiple classification scores between 0 and 1 as the probability of having a methylation profile related to every disease. To assess the sensitivity of the model, the testing cohort, which was not used for identification of the signature or construction of the SVM, was supplied to the model. To determine the specificity, we supplied all of the healthy subjects that were not used in the earlier stages of the study to the model. To understand whether this model is sensitive to other medical conditions representing developmental delay and intellectual disabilities, we tested, using the constructed model, a large number of subjects in our database with a confirmed clinical diagnosis of various diseases including autism spectrum disorders, imprinting defects, RASopathies, chromosomal aberrations, and Down syndrome. As well, we tested whether this classifier is sensitive to other diseases of epigenomic machinery for which no epi-signature was identified in this study. To further confirm that this classifier is not sensitive to the blood

cell type compositions, we downloaded normalized methylation data from isolated cell populations of healthy individuals from GEO (GSE35069)[16] and supplied them to our model for prediction.

Assessment of Ambiguous Case Subjects and Variants of Unknown Significance

The approved model was used to perform a prediction on the DNA methylation profiles of individuals with variants of unknown significance in the respective genes that were not previously included in the identification of the signature or in construction and validation of the classification model. In addition, a prediction was made on subjects with predicted sex discordance, samples obtained from tissues other than blood, healthy subjects carrying pathogenic mutations, and the single subject with Kabuki syndrome resulting from the less common *KDM6A* gene mutation.

Ethics Statement

This study has been approved by the Western University Research Ethics Boards (REB ID 106302) and the Hamilton Integrated Research Ethics Board (REB ID 13-653-T). All of the samples and records were de-identified before the study.

Results

Description of the Study Cohort

Figure 1 represents the flowchart of the study. Mutation analysis of the subjects with clinical features resembling each of the 14 syndromes (Table 1) identified a total of 285 subjects with pathogenic or likely pathogenic mutations, which were classified according to the American College of Medical Genetics (ACMG) guidelines. The remaining subjects (n = 176) carried benign variants or variants of unknown significance (VUS), leaving them unsolved. Our data also included healthy female carriers with pathogenic mutations in *KDM5C* (n = 8), fibroblast samples from individuals with Sotos syndrome (n = 3), and one individual affected with Kabuki syndrome with a pathogenic truncating mutation in *KDM6A*. Table S1 summarizes the mutation types and demographic characteristics where available. A sample of 196 randomly selected subjects from the 285 individuals with a pathogenic mutation was used as the discovery/training cohort for identification of epi-signatures as well as for training the classification model. The remaining 89 subjects were regarded as the testing cohort to be used for measuring the sensitivity of the classification model. For each disease group in the discovery/training cohort, a sex- and age-matched sample group four times larger (minimum n = 20) was selected from our reference healthy cohort (n = 650) for comparison. A total of 190 healthy samples never selected as a control for any of the diseases were later used to assess the specificity of the classification model. Table 1 shows the count, age, and sex distributions from every subject group in the discovery/training cohort along with the information from the matched control subjects.

Disease-Specific DNA Methylation Epi-signatures and Differentially Methylated Regions (DMRs)

The comparison of disease cohorts with their matched control subjects was performed using the subset of CpG probes that passed the quality assessment (Table 1). No reliable epi-signature was observed for five of the diseases tested here, including Saethre-Chotzen, Coffin-Siris, Coffin-Lowry, Rett, and Weaver syndromes, as the findings did not meet the criteria described in the methods. For other conditions (i.e., Floating-Harbor, ADCA-DN, Kabuki, ATRX, CHARGE, Sotos, GTPTS, SBBYSS, and Claes-Jensen syndromes), reliable epi-signatures were identified (Table 1). Of these, Sotos (n = 6,858) and ADCA-DN (n = 3,562) revealed the largest number of probes, mostly composed of hypomethylated CpGs. The identified probes from every cohort were confirmed to separate the subjects from the controls using hierarchical clustering (data not shown).

Consistent with the methylation profiles, the bump hunting approach did not identify any genomic segment to be differentially methylated in subjects with Saethre-Chotzen, Coffin-Siris, Coffin-Lowry, Rett, and Weaver syndromes. For the other nine syndromes, however, multiple genomic coordinates containing a minimum of 3 consecutive CpG probes, an average regional methylation difference > 0.10, and a family-wise error rate (FWER) < 0.01 were identified. Subjects with Sotos syndrome showed the largest number of identified regions (n = 1,372), mostly composed of hypomethylated segments (data not shown).

To further ensure that adjustment for blood cell type compositions has not masked a potential methylation profile in the five syndromes with negative results, we repeated the analysis without inclusion of the blood cell type estimates correction. Similar to the previous analyses, a significant methylation profile was not detected. At this stage, we concluded that either no epigenomic profile existed for these five conditions, or their methylation changes were too obscure to pass the thresholds and quality assessment criteria set in this study (see Material and Methods). Thus, our subsequent analyses described below focused on the syndromes for which a differential epigenomic profile was observed: i.e., Floating-Harbor, ADCA-DN, Kabuki, ATRX, GTPTS, SBBYSS, CHARGE, Sotos, and Claes-Jensen syndromes.

Overlap between the Epi-signatures and DMRs

Of the total number of 15,408 probes that composed the epi-signatures of the 9 conditions, only 1,598 (~10%) CpGs were shared across more than one disease group. Among these conditions, Floating-Harbor and Claes-Jensen syndromes were found to share the largest proportion of their epi-signatures with others (mainly with Sotos), whereas this number was minimal to negligible for other conditions, including ADCA-DN which, despite having a large number of CpGs in its epi-signature, shared less than 3% of its probes with other diseases (Figure 2). The

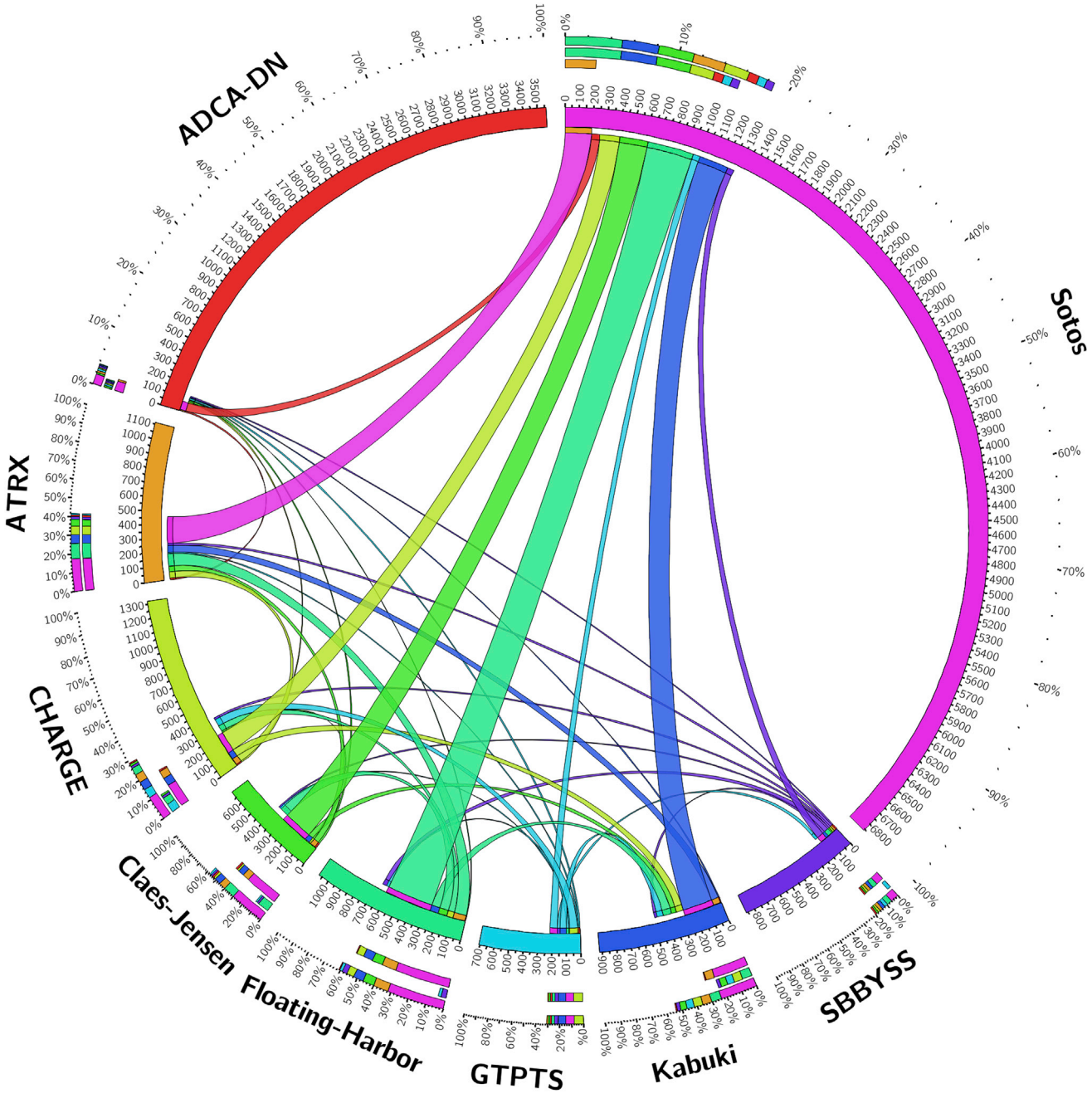

Figure 2. Quantity of Probes from the Epi-signatures of Every Nine Conditions that Are Shared with Each Other
Thickness of the bonds represents the number of the shared probes by every two diseases as shown by digits on the circumference of the plot. This plot does not visualize the 217 probes that are shared by more than two conditions.

number of probes that were shared by more than two conditions was limited to 217 probes. Table S2 shows the 217 probes and the related conditions. A gene ontology analysis of the 217 probes using missMethyl package found multiple ontology terms related to histone modifications to be enriched in the harboring genes, including S-methyltransferase activity and histone-lysine N-methyltransferase activity (Table S3). Next, genes overlapping these probes were evaluated using the ReactomePA package to identify pathways enriched in them (Table S4). Lysine histone transferase activity was found as the only enriched pathway with a multiple testing corrected p value < 0.05, which is composed of four genes (*PRDM9*, *SETDB1*, *HIST1H3E*, *NSD1*). These four genes also contained the most number of probes shared by different conditions (Table S2), which included *PRDM9* with five probes found in all of the conditions except for ADCA-DN, GTPTS, and SBBYSS; *HIST1H3E*, with two probes shared by Floating-Harbor, ATRX, Claes-Jensen, GTPTS, and SBBYSS syndromes; and *NSD1* and *SETDB1* with three and two probes, respectively, shared by Sotos, Floating-Harbor, and Claes-Jensen syndromes. Table S2 also shows that 21 probes in

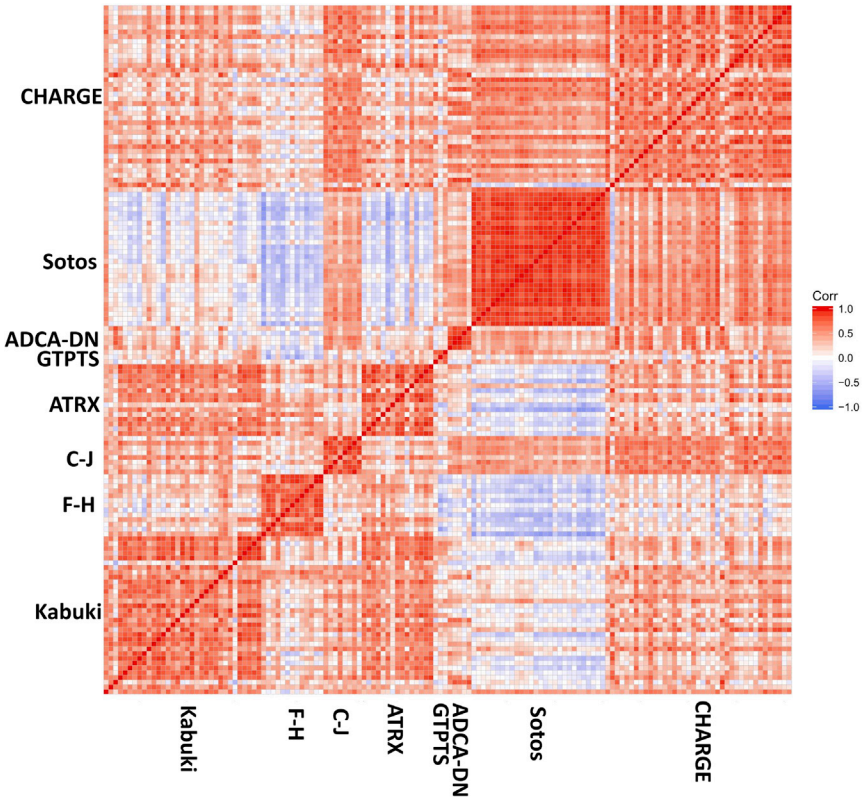

Figure 3. Pairwise Correlation between Samples with Different Conditions using the Methylation Values of Ten Probes that Are Shared by More than Three Conditions

Red represents positive and blue represents negative correlation. Every visible square represents the correlations of one subject on x axis with its correspondence on y axis. Abbreviations: F-H, Floating-Harbor; C-J, Cales-Jensen

HOXA5 (MIM: 142952) are shared by three conditions (Kabuki, CHARGE, and GTPTS).

A pairwise correlation analysis on the methylation levels of all of these 217 probes for each affected subject in the disease group revealed that while samples from the same condition showed the strongest correlation with each other, only a weak to moderate positive correlation existed among the samples from across disease groups. The only exception was with subjects with Floating-Harbor and Sotos syndromes, which revealed a moderate negative correlation with each other, despite sharing the most number of probes of all. We further narrowed down the probes to 18 CpGs that were shared by more than three disease groups and re-evaluated the correlation of the subjects based on the methylation levels of these 18 probes. We observed a greater degree of negative correlations between different syndrome groups (Figure 3). The members of every condition, however, correlated well with each other in both of these analyses.

Table 2 shows the genomic coordinates that are shared between every two diseases. Similar to the results described at single probe analysis, Floating-Harbor and Sotos syndromes shared the most differentially methylated coordinates (n = 30), more than half of which showed an opposite direction of change in methylation levels by the two syndromes. As previously observed by the shared CpGs (Table S2), a segment in the promoter of *PRDM9* was among the regions containing at least three consecutive hypomethylated probes in both ATRX and Sotos syndromes (Figure 4). The same segment contains disperse probes hypomethylated in Claes-Jensen syndrome but hypermethylated in Floating-Harbor syndrome (Table S2). The only region with multiple consecutive differentially methylated probes that was shared by more than two syndromes was a segment in the promoter of *HOXA5*. This segment is significantly hypomethylated in GTPTS but hypermethylated in both CHARGE and Kabuki syndromes (Table 2).

Overall, our analyses showed limited overlap between the epi-signatures, which tend to occur only in a few genomic coordinates. Within these regions, the methylation levels do not correlate well across the diseases, and in many cases the direction of methylation change (hypo- versus hypermethylation) is opposite. These results suggested that it might be possible to combine all of the epi-signatures for building a single classification model for concurrent classification of all of the diseases.

Development and Validation of a Classification Model for Prediction of Disease Classes

To develop a classification model, a multi-class support vector machine (SVM) with linear kernel was trained using a subset of 929 of the most differentiating and non-redundant probes selected from the epi-signature of every syndrome (Table S5). Due to small sample size for GTPTS (n = 3) and SBBYSS (n = 1), these conditions were not included in the model, and thus the training was performed only for the remaining seven syndromes. One model was trained for all of the affected subjects from these seven syndromes (n = 141) and the control subjects. Only the samples from the discovery/training cohort were used for training. The model was set to generate seven classification scores between 0 and 1 as the probability of having a methylation profile related to any of the seven syndromes. Ten-fold cross-validation of this model revealed an accuracy of 99.6%, and it correctly predicted the class of all of the 141 affected subject that were used for its training (Figures 5A–5G).

To determine the sensitivity of our model, 71 subjects from the testing cohort were supplied to the classification

Table 2. Overlapping Genomic Coordinates with Differential Methylation in Nine Conditions Compared to Control Subjects

Chr	Disease 1	Coordinates 1	Methylation Difference 1	Probes 1	Disease 2	Coordinates 2	Methylation Difference 2	Probes 2	Direction of Change	Overlapping Gene(s)	Distance to CpG Island (bp)
19	Floating-Harbor	8591364–8591776	−0.4	4	Kabuki	8591364–8591776	−0.28	4	same	*MYO1F*	0
11	Floating-Harbor	65360123–65360327	−0.27	3	Kabuki	65360123–65360327	−0.19	3	same	*KCNK7*	0
1	Sotos	227746111–227747468	−0.35	7	Kabuki	227746191–227746882	0.14	4	opposite	–	0
3	Sotos	49170496–49171051	−0.22	7	Kabuki	49170599–49170794	0.16	4	opposite	*LAMB2*	12,032
5	Sotos	1856325–1857828	−0.2	7	Kabuki	1856713–1857477	0.16	4	opposite	–	0
12	Sotos	133179887–133180698	−0.25	5	Kabuki	133179887–133180238	0.11	4	opposite	*LRCOL1*	0
5	Sotos	101119084–101119766	−0.24	5	Kabuki	101119084–101119566	0.11	4	opposite	–	512,283
7	Sotos	27170241–27170552	0.19	6	Kabuki	27170388–27170994	−0.14	13	opposite	*HOXA4*	0
19	Sotos	51330265–51330469	−0.24	4	Kabuki	51330265–51330469	0.13	4	opposite	*KLK15*	0
15	Sotos	29968032–29968195	−0.3	3	Kabuki	29968032–29968195	0.18	3	opposite	–	410
7	Sotos	27170717–27171051	0.15	9	Kabuki	27170388–27170994	−0.14	13	opposite	–	78
2	Sotos	109746691–109747003	−0.21	5	Kabuki	109746691–109746754	−0.14	4	same	*SH3RF3*	0
7	Sotos	142494148 - 142494492	−0.16	6	Kabuki	142494148–142494492	−0.15	6	same	–	71
5	Sotos	176827082–176827697	−0.16	5	Kabuki	176827392–176827697	−0.16	3	same	*PFN3*	0
10	Sotos	132099067–132100019	−0.17	3	Kabuki	132099067–132100019	0.15	3	opposite	–	109,734
7	CHARGE	27182493–27183946	0.23	20	Kabuki	27182493–27183816	0.16	18	same	*HOXA5, HOXA-AS3*	0
7	CHARGE	27184316–27184521	0.17	8	Kabuki	27184369–27184441	0.14	4	same	*HOXA-AS3*	0
17	Floating-Harbor	7486551–7486874	−0.24	7	Claes–Jensen	7486551–7486874	−0.29	7	same	–	0
1	Sotos	247694041–247694531	−0.25	6	Claes–Jensen	247694041–247694531	−0.24	6	same	*GCSAML,GCSAML-AS1, OR2C3*	0
5	Sotos	176559334–176559563	−0.36	3	Claes–Jensen	176559334–176559563	−0.32	3	same	–	0
13	Sotos	113242878–113243141	−0.21	3	Claes–Jensen	113242878–113243141	−0.32	3	same	*TUBGCP3 (396bp away)*	221

(*Continued on next page*)

Table 2. ***Continued***

Chr	Disease 1	Coordinates 1	Methylation Difference 1	Probes 1	Disease 2	Coordinates 2	Methylation Difference 2	Probes 2	Direction of Change	Overlapping Gene(s)	Distance to CpG Island (bp)
6	Sotos	32120625–32121433	−0.31	27	Floating-Harbor	32120773–32120933	−0.15	10	same	–	396
6	Sotos	31650735–31651249	−0.17	17	Floating-Harbor	31650735–31650835	−0.19	6	same	–	0
6	Sotos	31650735–31651249	−0.17	17	Floating-Harbor	31650916–31651278	−0.15	11	same	–	0
6	Sotos	32847702–32847845	−0.29	10	Floating-Harbor	32847702–32847845	0.18	10	opposite	–	0
17	Sotos	36997420–36997740	−0.38	7	Floating-Harbor	36997420–36997740	−0.23	7	same	*C17orf98*	0
8	Sotos	39171866–39172120	−0.31	7	Floating-Harbor	39172020–39172120	0.2	6	opposite	*ADAM5 (61bp away)*	206,442
10	Sotos	50649666–50650248	−0.39	5	Floating-Harbor	50649666–50650248	0.18	5	opposite	–	42,882
6	Sotos	33282867–33283055	−0.16	12	Floating-Harbor	33282867–33283184	−0.18	19	same	–	0
2	Sotos	129659018–129659946	−0.24	7	Floating-Harbor	129659316–129659946	0.19	6	opposite	–	0
5	Sotos	83016779–83017553	−0.25	6	Floating-Harbor	83017000–83017644	0.19	6	opposite	*HAPLN1*	341
4	Sotos	11369349–11370872	−0.18	8	Floating-Harbor	11370314–11370565	0.25	3	opposite	*MIR572*	0
8	Sotos	102235927–102236831	−0.23	6	Floating-Harbor	102236283–102236831	0.23	5	opposite	–	0
2	Sotos	165811816–165812159	−0.26	5	Floating-Harbor	165811816–165812159	0.22	5	opposite	*SLC38A11*	112,922
7	Sotos	92672812–92673176	−0.25	5	Floating-Harbor	92672812–92673176	0.2	5	opposite	–	0
4	Sotos	155702409–155703138	−0.18	6	Floating-Harbor	155702409–155703138	0.19	6	opposite	*RBM46*	0
1	Sotos	1003126–1003529	−0.25	4	Floating-Harbor	1003126–1003529	−0.29	4	same	–	0
3	Sotos	159557552–159558031	−0.24	4	Floating-Harbor	159557552–159558031	0.23	4	opposite	*SCHIP1, IQCJ-SCHIP1*	74,528
2	Sotos	164204628–164204915	−0.24	4	Floating-Harbor	164204752–164205343	0.31	6	opposite	–	0

(*Continued on next page*)

Table 2. ***Continued***

Chr	Disease 1	Coordinates 1	Methylation Difference 1	Probes 1	Disease 2	Coordinates 2	Methylation Difference 2	Probes 2	Direction of Change	Overlapping Gene(s)	Distance to CpG Island (bp)
22	Sotos	50737978–50738890	−0.22	4	Floating-Harbor	50737978–50738890	−0.29	4	same	*PLXNB2*	0
12	Sotos	56617576–56617737	−0.28	3	Floating-Harbor	56617576–56617737	−0.23	3	same	–	258
9	Sotos	139258524–139259074	−0.26	3	Floating-Harbor	139258524–139259074	−0.23	3	same	*CARD9*	0
19	Sotos	3480363–3480672	−0.21	5	Floating-Harbor	3480363–3480672	−0.2	5	same	*SMIM24*	1,110
19	Sotos	49222892–49223278	−0.2	5	Floating-Harbor	49222892–49223278	−0.28	5	same	*MAMSTR*	0
4	Sotos	165898666–165898848	−0.19	5	Floating-Harbor	165898666–165898848	0.19	5	opposite	*TRIM61*	20,219
1	Sotos	45278971–45279349	−0.21	4	Floating-Harbor	45278971–45279349	−0.23	4	same	*BTBD19*	0
19	Sotos	2428350–2428677	−0.19	4	Floating-Harbor	2428350–2429209	−0.2	6	same	*LMNB2*	0
19	Sotos	1063624–1064218	−0.19	3	Floating-Harbor	1063624–1064218	−0.24	3	same	*ABCA7*	0
4	Sotos	46126066–46126253	−0.18	3	Floating-Harbor	46126066–46126448	0.25	7	opposite	*GABRG1*	265,650
12	Sotos	47219737–47219958	−0.14	9	Floating-Harbor	47219626–47219920	0.18	9	opposite	*SLC38A4*	4,954
5	Sotos	110062384–110062837	−0.16	6	Floating-Harbor	110062384–110062837	0.23	6	opposite	*TMEM232*	11,768
11	Sotos	32449254–32449638	−0.17	3	ADCA-DN	32449163–32450692	0.24	11	opposite	*WT1*	0
7	CHARGE	27137922–27138712	−0.15	4	Sotos	27137922–27138396	−0.17	3	same	*HOTAIRM1*	1,185
7	CHARGE	53254947–53255065	0.15	3	Sotos	53254947–53255065	0.18	3	same	–	0
5	CHARGE	161178574–161178796	0.14	3	Sotos	161178574–161178796	0.17	3	same	–	203,189
5	ATRX	23507450–23507752	−0.26	10	Sotos	23507573–23507656	−0.22	5	same	*PRDM9*	1,682,751
15	ATRX	39871808–39872186	−0.26	6	Sotos	39871876–39872186	−0.16	5	same	–	341
3	ATRX	109056349–109056897	0.29	4	Sotos	109056349–109056897	−0.26	4	opposite	*DPPA4*	219,259
6	ATRX	34499314–34499504	−0.28	4	Sotos	34499314–34499504	−0.17	4	same	*PACSIN1*	0
16	ATRX	58019866–58019984	−0.37	3	Sotos	58019866–58019984	−0.27	3	same	*TEPP*	0

(*Continued on next page*)

Table 2. *Continued*

Chr	Disease 1	Coordinates 1	Methylation Difference 1	Probes 1	Disease 2	Coordinates 2	Methylation Difference 2	Probes 2	Direction of Change	Overlapping Gene(s)	Distance to CpG Island (bp)
10	ATRX	43698834–43699070	−0.33	3	Sotos	43698834–43699070	−0.32	3	same	*RASGEF1A*	656
18	ATRX	77905119–77905751	−0.18	10	Sotos	77905663–77905751	0.16	4	opposite	*PARD6G-AS1 (56bp away)*	0
1	ATRX	228291486–228291705	0.23	5	Sotos	228291591–228291682	0.18	3	same	–	301
18	ATRX	74499372–74499584	0.2	5	Sotos	74499372–74499584	−0.17	5	opposite	–	0
1	ATRX	3634867–3635100	−0.22	4	Sotos	3634867–3635100	−0.2	4	same	*TP73*	0
7	GTPTS	27182493–27185282	−0.36	51	Kabuki	27182493–27183816	0.15	18	opposite	*HOXA5*	0
7	GTPTS	27182493–27185282	−0.36	51	CHARGE	27182493–27183946	0.23	20	opposite	*HOXA5*	0
14	GTPTS	54418728–54418851	−0.38	3	SBBYSS	54418728–54418851	−0.2	3	same	*BMP4*	0

model. All of these subjects received a score of close to 1 for the syndrome to which they belonged to but a low score for the other six conditions, showing that our model is 100% sensitive for detecting any of the seven conditions examined (Figures 5A–5G). To estimate the specificity of our method, we used 190 normal samples from our reference cohort that were not used for epigenomic profiling of any of the diseases in this study or training the model. All of these samples received scores close to zero by our model for all of the seven disease groups, showing a specificity of 100% (Figure 5H).

To determine whether this classification algorithm is sensitive to the composition of blood cell types, we downloaded a total of 60 methylation array data from 6 healthy individuals,[16] each being assayed for whole blood, peripheral blood mononuclear cells (PBMCs), and granulocytes, as well as for seven isolated cell populations ($CD4^+$ T cells, $CD8^+$ T cells, $CD56^+$ NK cells, $CD19^+$ B cells, $CD14^+$ monocytes, neutrophils, and eosinophils). The methylation data from these 60 samples were imported into our classification model. Our algorithm assigned a score of close to zero for having any of the seven conditions to all of the files, suggesting that the composition of blood cells will not influence the performance of our model. The scores obtained in this assessment are presented in Table S6.

The 7-Disease Classification Model Is Not Sensitive to Other Developmental Conditions

Next, we examined whether this classification model could distinguish between the epigenomic profile of the seven conditions and the five other previously tested syndromes (Saethre-Chotzen, Coffin-Siris, Coffin-Lowry, Rett, and Weaver syndromes) for which we did not find an epi-signature. Hence, the model was supplied with all of the 69 samples with pathogenic mutations in genes from these five syndromes, as well as to 55 subjects with clinical features similar to these conditions but with VUS variants in the related genes. All of these subjects received low scores for any of the seven disease classes in the algorithm. Next, to determine the performance of this classifier in subjects with developmental delay/intellectual disability (DD/ID) of other etiologies, the model was supplied with the methylation levels of additional cohorts composed of individuals with autism spectrum disorders (n = 146), various chromosomal abnormalities (n = 12), Down syndrome (n = 7), imprinting conditions (Angelman, Beckwith-Wiedemann, Prader-Willi syndrome, n = 50), and various forms of RASopathies (n = 97). All of these subjects had a confirmed clinical and molecular diagnosis of their respected conditions. We found that similar to our previous observations, each of these subjects received low scores for all of the seven diseases in our classification model, further demonstrating that the epi-signatures of these seven conditions are highly specific. These results are demonstrated in Figure 5H.

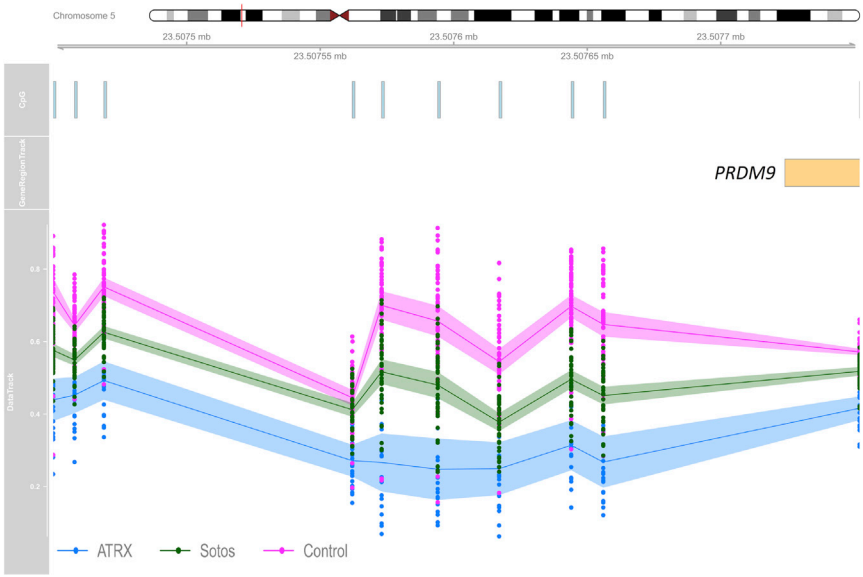

Figure 4. Hypomethylation of the 5′ UTR of *PRDM9* in Both Sotos-Affected and ATRX-Affected Subjects
The figure illustrates a 302-base pair region containing 10 CpG probes overlapping 5′ UTR of *PRDM9*. From top to bottom: chromosome ideogram, CpG probes, gene region, and methylation level data. Blue, ATRX; green, Sotos; pink, controls; line, average methylation; shadow, 95% confidence interval; dots, methylation values from every single sample (0-1).

Classification of Variants of Unknown Significance, Healthy Carriers, Atypical Case Subjects, and Non-blood Methylation Profile

We next applied our classification model to a heterogeneous group of subjects (n = 176) from our cohort displaying some clinical features related to multiple of the aforementioned developmental conditions for which the variant pathogenicity was uncertain (Table S7; Figure 6), in addition to a few other samples from sources other than blood, healthy carriers, and a subject with Kabuki syndrome resulting from *KDM6A* mutations. Of the 36 subjects within this cohort with clinical features of Kabuki syndrome but with benign or VUS variants in *KMT2D*, seven were predicted to have a DNA methylation profile specific to Kabuki syndrome. The only subject with Kabuki syndrome with a *KDM6A* mutation also received a high classification score for Kabuki syndrome. In this cohort, there was also one individual with Floating-Harbor syndrome with a discordance between the predicted and the reported sex. This subject was predicted to belong to the Floating-Harbor class by our model. Later assessment confirmed that it was only a sample-labeling error. Our data contained healthy female carriers of pathogenic mutations in *KDM5C* (n = 8). All of these subjects received low scores for any of the seven conditions, including the Claes-Jensen syndrome. Two subjects with similar clinical presentations to ATRX but with no confirmed variants in the *ATRX* gene were predicted as not having any of the seven diseases including ATRX. Half of the 16 samples with VUS in *NSD1* were predicted, based on the methylation score, as having ATRX. Our data contained four subjects suspected of having CHARGE syndrome for which sequence analysis results were not available at the time of the study, but who were all predicted to have CHARGE syndrome by our methylation prediction model. Subsequent sequence analysis reports of *CHD7* for these four subjects showed that truncating exon-intron boundary mutations exist in two of them (GenBank: NG_007009.1, c.4533+1G>A [p.(=)]; GenBank: NG_007009.1, c.6937−1G>A [p.(=)]). The other two subjects shared one splice site mutation (GenBank: NG_007009.1, c.5405−17G>A [p.(=)]) which has previously been reported in CHARGE syndrome and confirmed to induce aberrant splicing using RNA-seq analysis in another subject with CHARGE syndrome. Of the remaining 51 subjects with similar phenotypic presentations to CHARGE syndrome, for whom *CHD7* mutation screening had not provided a conclusive result, 23 were predicted as CHARGE, 27 were assigned low likelihood for all of the seven conditions, and 1 was predicted as having a methylation profile similar to Kabuki syndrome. Further follow-up and sequencing of this subject identified a pathogenic truncating variant (p.Arg5097* [c.15289C>T]; GenBank: NG_027827.1) in *KMT2D*, which confirmed the diagnosis of Kabuki syndrome. The three fibroblast samples from the Sotos-affected subjects were each predicted to have a similar methylation profile to the peripheral blood DNA epi-signature of Sotos syndrome. These findings recommended the high capacity of our method to assign new molecular diagnoses to subjects for whom the sequence analysis was not available or not interpretable, or the initial clinical suspicion was not correct. The complete classification outcome for these subjects is presented in Figure 6 and Table S7.

Discussion

Here we report syndrome-specific, minimally overlapping, DNA methylation epi-signatures in peripheral blood from a large cohort of individuals who possess pathogenic mutations in genes involved in the regulation of epigenomic machinery. Fourteen syndromes were represented in this cohort (Table 1), with reliable epi-signatures being observed for nine syndromes evaluated. We also describe here the development of a single classification algorithm for seven of these syndromes which has a highly sensitive and specific performance in predicting disease class, and which confidently rejects the probability of healthy individuals or other subjects with DD/ID to be affected by any of these seven conditions. Our results support the

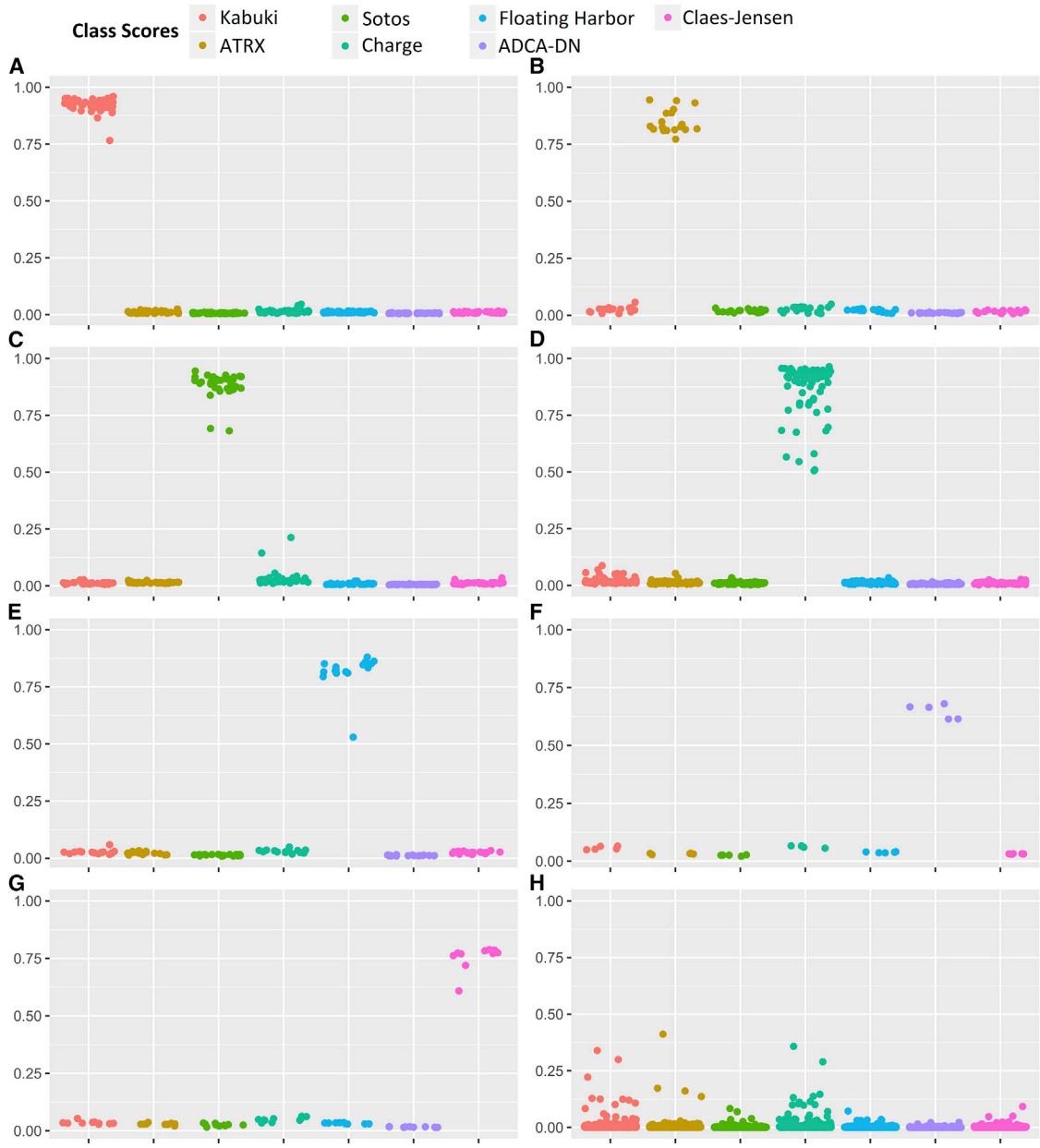

Figure 5. Probability Scores Generated by the Classification Model

A 7-disease SVM classifier concurrently generates seven scores for every subject as the probability of having a DNA methylation profile similar to any of the seven diseases with a confirmed DNA methylation signature. y axis represents scores 0-1, with higher scores indicating a higher chance of carrying a methylation profile related to any of the seven conditions. x axis represents the seven classification scores generated for the same group of tested subjects. These include the probability of having a similar DNA methylation profile to Kabuki syndrome, ATRX, Sotos syndrome, CHARGE syndrome, Floating-Harbor syndrome, ADCA-DN, and Claes-Jensen, respectively. By default, the SVM classifier defines a cut-off of 0.5 for predicting the class; however, the vast majority of the tested individuals received a score close to 0 or 1. Therefore, for the purpose of better visualization, the points are jittered. Every point represents the probability score received for a single sample. This figure represents scores obtained by both the subjects in the training and testing cohorts. Shown are probability scores for belonging to any of the seven classes for: 44 subjects with Kabuki syndrome (A); 19 subjects with ATRX syndrome (B), 38 subjects with Sotos syndrome (C), 79 subjects with CHARGE syndrome (D), 17 subjects with Floating-Harbor syndrome (E), 5 subjects with ADCA-DN (F), and 10 subjects with intellectual disability due to *KDM5C* (G). The last panel (H) shows the probabilities of belonging to any of the seven disease groups for 436 subjects with other conditions presenting with DD/ID including diseases of epigenomic machinery for which no epi-signature was found, multiple chromosomal aberrations, Down syndrome, various forms of RASopathies, autism spectrum disorders, and imprinting defect conditions, together with 190 healthy control subjects which were not used in any previous step in the study.

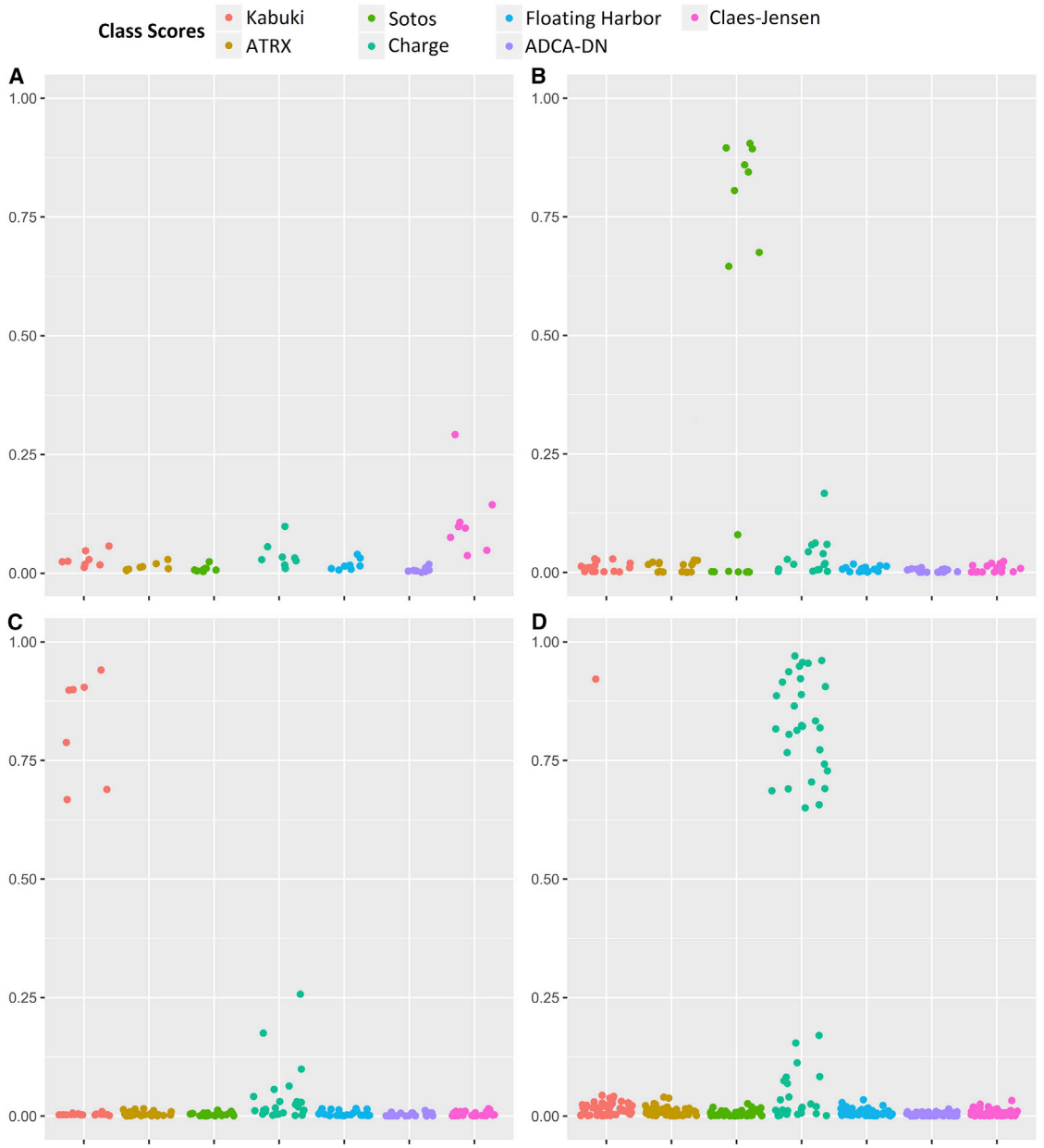

Figure 6. Probability Scores Generated by the Classification Model for Case Subjects with Uncertain Diagnosis Carrying Benign or VUS Variants and Healthy Carriers

A 7-disease SVM classifier concurrently generates seven scores for every subject as the probability of having a DNA methylation profile similar to any of the seven diseases with a confirmed DNA methylation signature. y axis represents scores 0-1, with higher scores indicating a higher chance of carrying a methylation profile related to any of the seven conditions. x axis represents the seven classification scores generated for the same group of tested subjects. These include the probability of having a similar DNA methylation profile to Kabuki syndrome, ATRX, Sotos syndrome, CHARGE syndrome, Floating-Harbor syndrome, ADCA-DN, and Claes-Jensen, respectively. Every point represents the probability score obtained for a single sample. Shown are probability scores for belonging to any of the seven classes for: 8 healthy female carriers with pathogenic mutations in *KDM5C* (A), 16 subjects with VUS variants in *NSD1* (B), 36 subjects with VUS and benign variants in *KMT2D* (C), and 55 subjects with similar features to CHARGE syndrome but no sequence data available or with VUS variants in *CHD7* (D).

hypothesis that the disruption of normal functions by the genes that regulate epigenetic marks, particularly those associated with histone modifications, generate unique DNA methylation epi-signatures with minimal overlaps across the conditions.

Epigenetic marks are established and maintained by an intricate network of proteins that shape the epigenome. These proteins function as writers that establish the epigenetic marks, readers that interpret them, and erasers that remove the epigenetic marks.[17] The fourth group of remodeller proteins comprises chromatin remodeling complexes that further regulate the epigenome to ensure the contextual (tissue-specific and temporal) accuracy of chromatin. Together, when properly integrated, these highly

coordinated and interactive proteins act within multi-protein, multifunctional complexes on chromatin to ensure the tissue- and lineage-specific transcriptional activation and suppression that is essential for proper embryonic and fetal development.[13] Mutations in the genes encoding these proteins can compromise the functional complexity and integration of the epigenetic machinery, lead to errors in the epigenetic signatures on chromatin, and result in a broad range of pediatric genetic disorders,[2,3] including those investigated in the present study. Thus, while the original genetic mutation in that particular epigenetic regulator may be primarily detrimental to normal neurodevelopment (and be a significant contributor to that clinical phenotype), all cells may carry an epigenetic echo (or legacy) of that original genetic defect in the form of a DNA methylation signature that may also contribute to the clinical phenotype. This association between a primary mutation in a gene encoding a regulator of chromatin accessibility and a secondary pattern of DNA methylation is likely due to the cross-talk between those interconnected pathways responsible for the post-translational histone modifications and DNA methylation.[17]

In the context of the cohort addressed in this paper, ADCA-DN and Rett syndrome resulted from mutations in proteins that write (DNMT1) and read (MeCP2) methylation marks on the DNA. Other syndromes resulted from errors in writers (KMT2D, NSD1, and KAT6B) and erasers (KDM5C) of the histone marks, and chromatin remodeling proteins (CHD7, SRCAP, ATRX). Our results show that syndromes caused by defects in chromatin remodeling proteins (CHARGE, Floating-Harbor, and ATRX), DNA methylation writer (ADCA-DN), histone mark erasers (Claes-Jensen syndrome), and histone mark writers (GTPTS, SBBYSS, Sotos, and Kabuki) generate specific DNA methylation signatures. On the other hand, in case subjects where the defect occurs in reading the methylation marks, such as in Rett syndrome, our study did not observe any consistent change in DNA methylation patterns, suggesting that the mutations in the multifunctional MeCP2 protein that lead to extensive temporal and spatial errors in levels of metabolites and gene expression across the entire genome, affect the function of multiple target pathways across cell lineages, rather than related to any lineage-specific epi-signature;[18] however, a possibility of a methylation signature in an alternate tissue cannot be discounted.

Overlap in the DNA Methylation Signatures Is Limited to the Initiating Event

Phenotypic overlap is a common finding in various conditions resulting from the disruption of the epigenomic machinery. Clinical overlap between CHARGE and Kabuki syndrome and the difficulty in distinguishing between the two conditions by facial features alone has been reported.[19] Sotos syndrome can be challenging to differentiate from Weaver or Beckwith-Wiedemann syndromes.[20] Similarly, Floating-Harbor syndrome exhibits some clinical overlap with Rubinstein-Taybi syndrome (MIM: 180849) (another disease of epigenomic machinery) as a result of the molecular interaction between the genes causing the two syndromes.[21] Bjornsson et al.[2] suggest that the phenotypic overlap that is seen in some imprinting disorders and a group of multiple congenital anomalies and DD/ID syndromes also result from mutations in the genes encoding components of the epigenetic machinery. Thus, overlap in these shared molecular targets and biological pathways may be the basis for phenotypic overlap in many of such conditions. The shared presence of a small number of genes known to be crucial to embryonic development within these DNA methylation profiles of the conditions we evaluated has not been thoroughly investigated. Butcher et al.[9] showed a gain of methylation in *HOXA5* in both Kabuki and CHARGE syndromes, which was also observed in this study (Table 2). Additionally, we observed the same region to be significantly hypomethylated in subjects with GTPTS. Butcher et al.[9] also report a segment in the *MYO1F* (MIM: 601480) gene body to be the most hypomethylated region in Kabuki syndrome-affected individuals. Our previous study on subjects with Floating-Harbor syndrome found the same region to be significantly hypomethylated.[5] Our current data also shows that two CpG probes from the same *MYO1F* region are also differentially methylated in CHARGE syndrome (Table S2), although with an opposite direction of change (hypermethylation) to those seen in Kabuki and Floating-Harbor syndromes. Hence, the primary mutational events in epigenetic regulators associated with these syndromes may initiate aberrant DNA methylation and then hardwire these methylation changes into the epigenome as abnormal repressive (or activation) signals at loci encoding developmentally relevant transcription factors, many of which are observed in our study (Table 2) including *WT1* (MIM: 607102) and *MAMSTR* (MIM: 610349) (Sotos), *DPPA4* (MIM: 614125) (ATRX), and *BMP4* (MIM: 112262) (GTPTS).

The concurrent screening of all of the single probes and regions with differential methylation levels in any of the diseases reported in this study have found that the existing overlap is limited to few specific regions and sites. Examination of DMRs, except for the segment mentioned above in *HOXA5*, did not reveal any region with a minimum of three consecutive probes to be shared by more than two conditions, and half of the shared regions (including *HOXA5*), showed an opposite direction of change in methylation levels. At the single CpG probe level, we found only 217 probes that were shared by more than two disease groups, the methylation levels of which were not strongly correlated in different conditions or were negatively correlated as observed between Sotos and Floating-Harbor, despite sharing the largest number of probes. Limiting the analysis to the probes shared by more than three conditions identified further discordance between the disorders (Figure 3). In addition, ADCA-DN shared the least number of probes and DMRs (only 1

DMR with opposite methylation change; Table 2) with other conditions, possibly due to the causative gene being a DNA methyltransferase versus other conditions resulting from histone modifications errors. Also, GTPTS and SBBYSS, which are caused by defects in the same gene, did not reveal any overlap in their epi-signatures and the total number of the probes they did share was less than a few dozen.

These limited shared probes were enriched and most frequently occurred in the genes related to laminin interaction (*LAMB2/LAMC1*), non-integrin membrane-ECM interactions (*LAMB2/LAMC1/NRXN1*), and synapse function (*PTPRF/DLG4/NRXN1*). Also, an intriguing finding was that multiple probes were shared across syndromes and were associated with genes functioning in pathways involving histone acetylation and methylation. Multiple ontology terms related to histone modifications were found to be enriched in the genes harboring these signatures, including S-methyltransferase activity and histone-lysine N-methyltransferase activities (Tables S3 and S4). These pathways involved three genes possessing histone methyltransferase activity (*PRDM9*, *SETDB1*, *NSD1*) and one coding for a subunit of histone H3 (*HIST1H3E*).

These results indicate that the shared pathways across these developmental syndromes include the initiating events leading to the disruption of the histone modification machinery. Once the initial mutation is established, each condition encompasses differential downstream paths leading to the generation of distinct epi-signatures. This is further supported by the single subject in our database with *KDM6A* mutation, whom we had predicted as having Kabuki syndrome, using the same model that is built on the epigenetic signature of *KMT2D* mutations. The immediate targets of these two genes are not overlapping; however, they mutually regulate a large number of genes downstream[22] and their disruption leads to a single medical condition with indistinguishable clinical features. This further suggests that the epigenetic profile of these conditions is additionally composed of the downstream changes unique to the syndrome rather than the primary epigenomic event. In addition to the aforementioned components related to histone acetylation and methylation, the *HOXA5* promoter was found to be differentially methylated in three syndromes, although with an opposite direction of methylation change. *HOXA5* encodes a transcription factor that spatiotemporally regulates the body segmentation and morphogenesis during development. It is well recognized that its expression is tightly regulated by the methylation status of its promoter.[23] Hypermethylation of *HOXA5* promoter in CHARGE and Kabuki and its hypomethylation in GTPTS may partly explain the similarity in the phenotypic overlap between CHARGE and Kabuki syndromes, but not in GTPTS, supporting the concept that different epigenetic changes in a region might result in alternative phenotypic outcomes.

A Model Accounting for DNA Methylation Signatures as Legacies of Mutations in Genes of the Epigenetic Machinery

From our results, we propose a mechanistic model by which syndrome-specific DNA methylation signatures arise and are maintained. Central to our model is the concept that the initial dysfunction during early development, which is linked to a specific mutated epigenetic reader, writer, eraser, or remodeler, leaves a specific clinically identifiable and syndrome-related phenotype. In addition, a broader methylation legacy or "echo" remains throughout the genome. That epigenetic echo, reflected in (but not necessarily limited to) DNA methylation, may be a secondary contributor to the overall clinical phenotype through inappropriate activation (or repression) of gene expression in a cell-lineage-dependent manner. For example, Kabuki is among a number of syndromes (including CHARGE and Rubinstein-Taybi) that in addition to displaying clinical features of intellectual disability and distinct craniofacial features, can also present with immune dysfunction. This clinical feature can increase susceptibility to infections and an inability to generate or maintain immunological memory.[24] Similarly, altered *KMT2D* function appears to impact cardiac development and cardiomyocyte function and may lead to defects in ion transport.[24] Several of the target genes reported in these papers and associated with these physiological functions were detected in the epi-signatures reported here in our paper, supporting the efficacy and surrogacy of peripheral blood DNA as a source for identifying the Kabuki-specific epi-signature. Furthermore, our approach raises the potential to directly identify epigenetically silenced genes that could be the focus of targeted epigenetic or gene-editing therapies. Such therapies could be focused either on the primary mutation harbored by the epigenetic machinery gene[25] or at the secondary targets that reside within the epi-signatures and display altered methylation within promoter or enhancer regions.

An additional application of the epigenetic signature detection extends beyond developmental syndromes to novel, targeted cancer therapy. In infrequent cases, young subjects with developmental syndromes including Kabuki, Coffin-Siris, and Sotos syndromes have been diagnosed with rare tumors at an early age.[26–28] This is not surprising given that these very same epigenetic regulatory genes (*KMT2D*, *NSD1*, and *SMARCB1*) are frequently mutated in many somatic cancer types.[17] Hence, such epigenetic information reflected by these epi-signatures has the potential to lead to the development of new targeted therapies for various adult cancers that are similarly associated with mutated genes encoding the epigenetic machinery.

DNA Methylation Signatures in Gene Defects across Different Tissues, Mutations, and Zygosity

A clinically applicable aspect of this study relates to whether DNA methylation signatures sourced from peripheral blood samples can act as a surrogate for aberrant

DNA methylation patterns that have their primary functional consequences in the developing nervous system and brain. Previous studies have generated conflicting results on the correlation of DNA methylation patterns across tissues and cell types.[29–32] We were able to directly address this issue of tissue surrogacy, in that training our prediction algorithm on the DNA methylation signatures from peripheral blood of Sotos syndrome-affected individuals also assigned high scores for the three fibroblast DNA samples also obtained from Sotos syndrome-affected individuals. This suggests that easily accessible peripheral blood DNA can be functionally comparable to clinically relevant yet inaccessible target tissues such as the brain. Thus, from a developmental standpoint, some DNA methylation changes associated with primary defects in epigenetic machinery genes are well established before the differentiation of organs in embryonic development and are maintained in these particular cell lineages or have similar functions across tissues.

In any event, this has to be confirmed by studying multiple organs of the affected individuals in different disorders. The limited findings that we have presented in this regard comply with this observation. While ADCA-DN occurs due to one mutation present in all organs, the disease exclusively exhibits features of the nervous system involvement. Bjornsson et al.,[2] based on this observation, postulates that various tissues show a specific and dosage-dependent response to disruption of epigenomic machinery. One observation in our study that may support this model is that the predictive algorithm that we trained using the epi-signature of the Claes-Jensen syndrome cannot detect any of the eight healthy female carriers of *KDM5C* pathogenic mutations. This indicates that the disruption of epigenetic machinery will be well represented in all tissues, but will lead to a functional disruption only when the dosage is beyond a critical point. This supports the clinical utility of peripheral blood as a surrogate tissue for screening for these conditions. Our study also suggests that the observed signature can be specific to the mutation type. GTPTS and SBBYSS are both caused by mutations in *KAT6B*. Depending on the location of the coding sequence and type of the mutation (i.e., missense versus frameshift), the gene product can be absent (loss of function) or abnormal (change of function), resulting in two distinct syndromes with overlapping features.[33] Our data show that these two conditions, despite resulting from mutations in the same gene, have distinct epi-signatures. This finding suggests that epi-signatures can potentially be useful in differentiating distinct diseases with overlapping features that are caused by a shared gene defect, where the gene sequencing alone cannot always assign a diagnosis.

The Clinical Implications of the Multi-class DNA Methylation Classification Model

The minimal overlap in the DNA methylation signatures of each of the seven conditions has allowed for training a classifier with the capacity of concurrent assessment of any given subject for all of the seven disease groups with a complete accuracy, sensitivity, and specificity. Furthermore, using datasets generated across multiple institutions, we have demonstrated generalized applicability of this approach in clinical setting that is not sensitive to the confounders such as batch effect. The model not only distinguished related disorders of the epigenetic machinery from each other, but also assigned low probability scores to subjects with various forms of DD/ID ranging from autism spectrum disorders to chromosomal aberrations and imprinting defect conditions that are not caused by mutations in these genes. This method will have a practical value in the clinical diagnosis of such genetic conditions where the disease is rare, clinical features are overlapping, and sequence mutation screening does not always produce conclusive results. In Kabuki syndrome, for instance, the mutation screening of *KDM6A* and *KMT2D*, together, identifies pathogenic mutations in only 70% of subjects.[34] A similar figure of 65%–70% is reported for atypical cases of CHARGE syndrome.[35] The vast majority of rare missense and in-frame in/del variants are currently not interpretable, and as VUSs present a significant challenge in clinical interpretation in diagnostic laboratories and genetic clinics. Some subjects also may carry pathogenic mutations in noncoding regions which are often not screened (e.g., promoter, intronic regions).

The method presented here provides an innovative approach to assigning pathogenicity to these variants by using the epi-signatures of each condition studied. Although the findings would have to be complemented with further clinical assessment, our model has shown its capacity to resolve a large proportion of the unsolved cases in our dataset by assigning new classifications to more than half of the subjects for whom sequence variant assessment alone had not provided a definitive answer. Of interest was a 4-year-old male with cleft palate, dysmorphic ears, microcephaly, and developmental delay, who was suspected of having CHARGE syndrome and who was screened for mutations in *CHD7* where only a heterozygous missense VUS variant (GenBank: NM_017780.3, c.2185A>G; GenBank: NM_017780.3, p.Lys729Glu) was found. The classification model we have described above predicted that this subject had an epigenomic profile similar, not to CHARGE syndrome (score = 0.02), but to Kabuki syndrome (score = 0.92). This subject was later confirmed to carry a pathogenic truncating variant in *KMT2D* and was subsequently diagnosed with Kabuki syndrome. It is worthwhile to note that in many of these conditions, such as Kabuki syndrome, the disease is often caused by *de novo* variants. Identification of a *de novo* variant increases the chance of the mutation being pathogenic, although this is not always sufficient to assign pathogenicity to a variant. Unlike the assessment of the "*de novo*" status of a genetic variant, DNA methylation analysis does not require parental sample DNA to assess the impact of the variant in the proband. However, where both analyses are available, combined evidence could be used to

further strengthen the case for variant pathogenicity status.

With more samples becoming available from such conditions and more diseases being studied for the discovery of epi-signatures, the methodology that we have presented here can provide the basis of application in routine molecular diagnostics. Microarray technology has been shown to be an invaluable tool for the diagnosis of epigenetic conditions including imprinting disorders,[36] and as we have shown in the present study, a combination of microarray technology with machine learning will have a great, and as yet relatively unexplored, potential in resolving a large number of subjects with unsolved etiology that are frequently seen in medical genetics practice.

Supplemental Data

Supplemental Data include seven tables and can be found with this article online at https://doi.org/10.1016/j.ajhg.2017.12.008.

Acknowledgments

We thank the families, the molecular genetics diagnostic laboratory at the Greenwood Genetic Center for identifying the mutations in the subjects, and the clinical geneticists at the Greenwood Genetic Center for making the clinical diagnosis. We also thank the Care4Rare Canada Consortium for the identification and molecular analyses of subjects with a subset of the neurodevelopmental conditions that were part of this study. E.A.-E. was supported by Children's Health Research Institute Epigenetics Trainee Award, funded by the Children's Health Foundation, London, Ontario, Canada. Dedicated to the memory of Ethan Francis Schwartz, 1996–1998.

Received: October 16, 2017
Accepted: December 10, 2017
Published: January 4, 2018

Web Resources

GenBank, https://www.ncbi.nlm.nih.gov/genbank/
GEO, http://www.ncbi.nlm.nih.gov/geo/
OMIM, http://www.omim.org/

References

1. Fahrner, J.A., and Bjornsson, H.T. (2014). Mendelian disorders of the epigenetic machinery: tipping the balance of chromatin states. Annu. Rev. Genomics Hum. Genet. *15*, 269–293.
2. Bjornsson, H.T. (2015). The Mendelian disorders of the epigenetic machinery. Genome Res. *25*, 1473–1481.
3. Berdasco, M., and Esteller, M. (2013). Genetic syndromes caused by mutations in epigenetic genes. Hum. Genet. *132*, 359–383.
4. Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. Nat. Rev. Genet. *10*, 295–304.
5. Hood, R.L., Schenkel, L.C., Nikkel, S.M., Ainsworth, P.J., Pare, G., Boycott, K.M., Bulman, D.E., and Sadikovic, B. (2016). The defining DNA methylation signature of Floating-Harbor syndrome. Sci. Rep. *6*, 38803.
6. Kernohan, K.D., Cigana Schenkel, L., Huang, L., Smith, A., Pare, G., Ainsworth, P., Boycott, K.M., Warman-Chardon, J., Sadikovic, B.; and Care4Rare Canada Consortium (2016). Identification of a methylation profile for DNMT1-associated autosomal dominant cerebellar ataxia, deafness, and narcolepsy. Clin. Epigenetics *8*, 91.
7. Schenkel, L.C., Kernohan, K.D., McBride, A., Reina, D., Hodge, A., Ainsworth, P.J., Rodenhiser, D.I., Pare, G., Bérubé, N.G., Skinner, C., et al. (2017). Identification of epigenetic signature associated with alpha thalassemia/mental retardation X-linked syndrome. Epigenetics Chromatin *10*, 10.
8. Choufani, S., Cytrynbaum, C., Chung, B.H.Y., Turinsky, A.L., Grafodatskaya, D., Chen, Y.A., Cohen, A.S.A., Dupuis, L., Butcher, D.T., Siu, M.T., et al. (2015). NSD1 mutations generate a genome-wide DNA methylation signature. Nat. Commun. *6*, 10207.
9. Butcher, D.T., Cytrynbaum, C., Turinsky, A.L., Siu, M.T., Inbar-Feigenberg, M., Mendoza-Londono, R., Chitayat, D., Walker, S., Machado, J., Caluseriu, O., et al. (2017). CHARGE and Kabuki syndromes: gene-specific DNA methylation signatures identify epigenetic mechanisms linking these clinically overlapping conditions. Am. J. Hum. Genet. *100*, 773–788.
10. Aref-Eshghi, E., Schenkel, L.C., Lin, H., Skinner, C., Ainsworth, P., Paré, G., Rodenhiser, D., Schwartz, C., and Sadikovic, B. (2017). The defining DNA methylation signature of Kabuki syndrome enables functional assessment of genetic variants of unknown clinical significance. Epigenetics *12*, 923–933.
11. Hamamori, Y., Sartorelli, V., Ogryzko, V., Puri, P.L., Wu, H.Y., Wang, J.Y., Nakatani, Y., and Kedes, L. (1999). Regulation of histone acetyltransferases p300 and PCAF by the bHLH protein twist and adenoviral oncoprotein E1A. Cell *96*, 405–413.
12. Sassone-Corsi, P., Mizzen, C.A., Cheung, P., Crosio, C., Monaco, L., Jacquot, S., Hanauer, A., and Allis, C.D. (1999). Requirement of Rsk-2 for epidermal growth factor-activated phosphorylation of histone H3. Science *285*, 886–891.
13. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al.; ACMG Laboratory Quality Assurance Committee (2015). 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. *17*, 405–424.
14. Schenkel, L.C., Schwartz, C., Skinner, C., Rodenhiser, D.I., Ainsworth, P.J., Pare, G., and Sadikovic, B. (2016). Clinical validation of fragile X syndrome screening by DNA methylation array. J. Mol. Diagn. *18*, 834–841.
15. Jaffe, A.E., Murakami, P., Lee, H., Leek, J.T., Fallin, M.D., Feinberg, A.P., and Irizarry, R.A. (2012). Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. Int. J. Epidemiol. *41*, 200–209.
16. Reinius, L.E., Acevedo, N., Joerink, M., Pershagen, G., Dahlén, S.E., Greco, D., Söderhäll, C., Scheynius, A., and Kere, J. (2012). Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PLoS ONE *7*, e41361.
17. Shen, H., and Laird, P.W. (2013). Interplay between the cancer genome and epigenome. Cell *153*, 38–55.
18. Ehrhart, F., Coort, S.L., Cirillo, E., Smeets, E., Evelo, C.T., and Curfs, L.M. (2016). Rett syndrome - biological pathways

leading from MECP2 to disorder phenotypes. Orphanet J. Rare Dis. *11*, 158.
19. Patel, N., and Alkuraya, F.S. (2015). Overlap between CHARGE and Kabuki syndromes: more than an interesting clinical observation? Am. J. Med. Genet. A. *167A*, 259–260.
20. Baujat, G., Rio, M., Rossignol, S., Sanlaville, D., Lyonnet, S., Le Merrer, M., Munnich, A., Gicquel, C., Colleaux, L., and Cormier-Daire, V. (2005). Clinical and molecular overlap in overgrowth syndromes. Am. J. Med. Genet. C. Semin. Med. Genet. *137C*, 4–11.
21. Hood, R.L., Lines, M.A., Nikkel, S.M., Schwartzentruber, J., Beaulieu, C., Nowaczyk, M.J., Allanson, J., Kim, C.A., Wieczorek, D., Moilanen, J.S., et al.; FORGE Canada Consortium (2012). Mutations in SRCAP, encoding SNF2-related CREBBP activator protein, cause Floating-Harbor syndrome. Am. J. Hum. Genet. *90*, 308–313.
22. Kim, J.H., Sharma, A., Dhar, S.S., Lee, S.H., Gu, B., Chan, C.H., Lin, H.K., and Lee, M.G. (2014). UTX and MLL4 coordinately regulate transcriptional programs for cell proliferation and invasiveness in breast cancer cells. Cancer Res. *74*, 1705–1717.
23. Watson, R.E., Curtin, G.M., Hellmann, G.M., Doolittle, D.J., and Goodman, J.I. (2004). Increased DNA methylation in the HoxA5 promoter region correlates with decreased expression of the gene during tumor promotion. Mol. Carcinog. *41*, 54–66.
24. Stagi, S., Gulino, A.V., Lapi, E., and Rigante, D. (2016). Epigenetic control of the immune system: a lesson from Kabuki syndrome. Immunol. Res. *64*, 345–359.
25. Harmeyer, K.M., Facompre, N.D., Herlyn, M., and Basu, D. (2017). JARID1 histone demethylases: emerging targets in cancer. Trends. Cancer *3*, 713–725.
26. Gossai, N., Biegel, J.A., Messiaen, L., Berry, S.A., and Moertel, C.L. (2015). Report of a patient with a constitutional missense mutation in SMARCB1, Coffin-Siris phenotype, and schwannomatosis. Am. J. Med. Genet. A. *167A*, 3186–3191.
27. Kulkarni, K., Stobart, K., and Noga, M. (2013). A case of Sotos syndrome with neuroblastoma. J. Pediatr. Hematol. Oncol. *35*, 238–239.
28. Karagianni, P., Lambropoulos, V., Stergidou, D., Fryssira, H., Chatziioannidis, I., and Spyridakis, I. (2016). Recurrent giant cell fibroblastoma: Malignancy predisposition in Kabuki syndrome revisited. Am. J. Med. Genet. A. *170A*, 1333–1338.
29. Ma, B., Wilker, E.H., Willis-Owen, S.A., Byun, H.M., Wong, K.C., Motta, V., Baccarelli, A.A., Schwartz, J., Cookson, W.O., Khabbaz, K., et al. (2014). Predicting DNA methylation level across human tissues. Nucleic Acids Res. *42*, 3515–3528.
30. Montaño, C.M., Irizarry, R.A., Kaufmann, W.E., Talbot, K., Gur, R.E., Feinberg, A.P., and Taub, M.A. (2013). Measuring cell-type specific differential methylation in human brain tissue. Genome Biol. *14*, R94.
31. Byun, H.M., Siegmund, K.D., Pan, F., Weisenberger, D.J., Kanel, G., Laird, P.W., and Yang, A.S. (2009). Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. Hum. Mol. Genet. *18*, 4808–4817.
32. Barault, L., Ellsworth, R.E., Harris, H.R., Valente, A.L., Shriver, C.D., and Michels, K.B. (2013). Leukocyte DNA as surrogate for the evaluation of imprinted loci methylation in mammary tissue DNA. PLoS ONE *8*, e55896.
33. Lonardo, F., Lonardo, M.S., Acquaviva, F., Della Monica, M., Scarano, F., and Scarano, G. (2017). Say-Barber-Biesecker-Young-Simpson syndrome and Genitopatellar syndrome: lumping or splitting? Clin. Genet. https://doi.org/10.1111/cge.13127.
34. Bögershausen, N., Gatinois, V., Riehmer, V., Kayserili, H., Becker, J., Thoenes, M., Simsek-Kiper, P.Ö., Barat-Houari, M., Elcioglu, N.H., Wieczorek, D., et al. (2016). Mutation update for Kabuki syndrome genes KMT2D and KDM6A and further delineation of X-linked Kabuki syndrome subtype 2. Hum. Mutat. *37*, 847–864.
35. Zentner, G.E., Layman, W.S., Martin, D.M., and Scacheri, P.C. (2010). Molecular and phenotypic aspects of CHD7 mutation in CHARGE syndrome. Am. J. Med. Genet. A. *152A*, 674–686.
36. Aref-Eshghi, E., Schenkel, L.C., Lin, H., Skinner, C., Ainsworth, P., Paré, G., Siu, V., Rodenhiser, D., Schwartz, C., and Sadikovic, B. (2017). Clinical validation of a genome-wide DNA methylation assay for molecular diagnosis of imprinting disorders. J. Mol. Diagn. *19*, 848–856.

The American Journal of Human Genetics, Volume *102*

Supplemental Data

Genomic DNA Methylation Signatures Enable Concurrent Diagnosis and Clinical Genetic Variant Classification in Neurodevelopmental Syndromes

Erfan Aref-Eshghi, David I. Rodenhiser, Laila C. Schenkel, Hanxin Lin, Cindy Skinner, Peter Ainsworth, Guillaume Paré, Rebecca L. Hood, Dennis E. Bulman, Kristin D. Kernohan, Care4Rare Canada Consortium, Kym M. Boycott, Philippe M. Campeau, Charles Schwartz, and Bekim Sadikovic

Table S1 - Characteristics of the patient cohort used in the study

Disease	**Id**	**Sex**	**Age**	**Gene**	**Mutation**
ATRX	MS0564	m	6.3	*ATRX*	c.6254G>A; p.R2085H
ATRX	MS0565	m	18	*ATRX*	c.736T>C; p.R246C
ATRX	MS0566	m	1.4	*ATRX*	c.6593A>G; p.H2198R
ATRX	MS0568	m	18.5	*ATRX*	c.758T>C; p.L253S
ATRX	MS0569	m		*ATRX*	c.109C>T; p.R37X
ATRX	MS0570	m	21	*ATRX*	c.4817G>A; p.S1606N
ATRX	MS0571	m		*ATRX*	c.865T>C; p.C220R
ATRX	MS0572	m	0.7	*ATRX*	c.5786A>G; p.K1929R
ATRX	MS0573	m	18	*ATRX*	c.952G>T; p.G249C
ATRX	MS0574	m	14	*ATRX*	c.730A>C; p.I244L
ATRX	MS0589	m	4.6	*ATRX*	c.7156C>T; p.R2386X
ATRX	MS0591	m	4.6	*ATRX*	c.536A>G, p.N179S
ATRX	MS0593	m	4.4	*ATRX*	deletion of exon 207
ATRX	MS0594	m		*ATRX*	c.109C>T; p.R37X
ATRX	MS0595	m	27	*ATRX*	c.7366_7367 InsA; p.M2456N fs X42
ATRX	MS0596	m	14.5	*ATRX*	c.109C>T; p.R37X
ATRX	MS0597	m	2.5	*ATRX*	c.736C>T; P.R246C
ATRX	MS0598	m	17.5	*ATRX*	c.109C>T; p.R37X
ATRX	MS0599	m	14	*ATRX*	c.109C>T; p.R37X
CHARGE	CHD7-1	f	4	*CHD7*	c.7282C>T (p.Arg2428*)
CHARGE	CHD7-10	f	0.02	*CHD7*	c.2905_2906del (p.Arg969Glyfs*25)
CHARGE	CHD7-11	m	12	*CHD7*	c.7636G>T (p.Glu2546*)
CHARGE	CHD7-12	m	19	*CHD7*	c.361delC (p.Gly121Valfs*90)
CHARGE	CHD7-13	m	19	*CHD7*	c.2504_2508delATCTT (p.Arg835Serfs*14)
CHARGE	CHD7-14	m	4	*CHD7*	c.7717-7720del (p.Gln2537*)
CHARGE	CHD7-15	m	11	*CHD7*	c.5458C>T (p.Arg1820*)
CHARGE	CHD7-16	m	7	*CHD7*	exon 1 deletion

CHARGE	CHD7-17	m	15	*CHD7*	c.5405-17G>A
CHARGE	CHD7-18	f	10	*CHD7*	c.5405-7G>A
CHARGE	CHD7-19	f	4	*CHD7*	c.2097-1G>A
CHARGE	CHD7-2	m	6	*CHD7*	c.3526C>T (p.Gln1176*)
CHARGE	CHD7-3	f	14	*CHD7*	c.934C>T (p.Arg312*)
CHARGE	CHD7-34	m		*CHD7*	c.5097dupA (p.Ala1700Serfs*37)
CHARGE	CHD7-36	m		*CHD7*	c.799G>T (p.Glu267*)
CHARGE	CHD7-4	m	13	*CHD7*	c562C>T (p.Gly188*)
CHARGE	CHD7-40	m		*CHD7*	c.1312C>T (p.Gln438*)
CHARGE	CHD7-41	m		*CHD7*	c.6322G>A (p.Gly2108Arg)
CHARGE	CHD7-46	f		*CHD7*	c.5050G>A (p.Gly1684Ser)
CHARGE	CHD7-47	f		*CHD7*	c.5210+3A>G
CHARGE	CHD7-5	m	18	*CHD7*	c.1327delATGGG (p.Met443Asnfs*130)
CHARGE	CHD7-51	f		*CHD7*	c.4087delC (p.Leu1363Serfs*9)
CHARGE	CHD7-52	m		*CHD7*	c.2498+1G>T
CHARGE	CHD7-53	f		*CHD7*	c.1918delG (p.Gly640Lysfs*71)
CHARGE	CHD7-55	f		*CHD7*	c.604 C>T (p.Gln202*)
CHARGE	CHD7-56	m		*CHD7*	c.4393C>T (p.Arg1465*)
CHARGE	CHD7-58	f		*CHD7*	c.5029C>T (p.Arg1677*)
CHARGE	CHD7-6	m	24	*CHD7*	c.2504_2508delATCTT (p.Tyr835Serfs*14)
CHARGE	CHD7-60	m		*CHD7*	c.3177T>G (p.Tyr1059*)
CHARGE	CHD7-61	m		*CHD7*	c.2839C>T (p.Arg947*)
CHARGE	CHD7-63	m		*CHD7*	c.4361_4362delAG (p.Gln1454Profs*21)
CHARGE	CHD7-64	f		*CHD7*	c.2362C>T (p.Gln788*)
CHARGE	CHD7-68	f		*CHD7*	c.8507delC (p.Pro2836Argfs*53)
CHARGE	CHD7-69	f		*CHD7*	c.3655C>T (p.Arg1219*)
CHARGE	CHD7-7	m	3	*CHD7*	c.1990G>T (p.Glu664*)
CHARGE	CHD7-71	f		*CHD7*	c.1141_1142delAT (p.Met381Alafs*23)
CHARGE	CHD7-72	m		*CHD7*	c.3082A>G (p.Ile1028Val)
CHARGE	CHD7-8	f	6	*CHD7*	c.3377dupT (p.Leu1126Phefs*46)

CHARGE	CHD7-9	m	9	*CHD7*	c.2585delA (p.Leu862Serfs*26)
CHARGE	MS0979	f	0.5	*CHD7*	c.7252C>T, p.R2418X
CHARGE	MS0982	m	0.8	*CHD7*	c.3655C>T, p.R1219X
CHARGE	MS0983	f	0.1	*CHD7*	c.7282C>T, p.R2428X
CHARGE	MS0984	f	0.005	*CHD7*	c.959_960delAG
CHARGE	MS0985	m	0.2	*CHD7*	c.5607+1G>A
CHARGE	MS0987	m	0.001	*CHD7*	c.4517dupG
CHARGE	MS0989	f	1.1	*CHD7*	c.6716delA
CHARGE	MS0991	f	15.4	*CHD7*	c.3654C>G, p.Y1218X
CHARGE	MS0992	m	2.1	*CHD7*	c.3640C>T, p.Q1214X
CHARGE	MS0995	m	0.1	*CHD7*	c.1972G>T, p.E658X
CHARGE	MS0996	f	3.2	*CHD7*	c.3778+1G>A
CHARGE	MS1002	m	13.5	*CHD7*	c.2464G>T, p.E822X
CHARGE	MS1003	f	0.8	*CHD7*	c.7957C>T, p.R2653X
CHARGE	MS1004	m	24.7	*CHD7*	c.4393C>T, p.R1465X
CHARGE	MS1005	f	0.01	*CHD7*	c.1480C>T, p.R494X
CHARGE	MS1007	m	0.003	*CHD7*	c.559C>T, p.Q187X
CHARGE	MS1011	f	1.5	*CHD7*	c.4015C>T, p.R1339X
CHARGE	MS1012	m	0.01	*CHD7*	c.2049delG
CHARGE	MS1013	f	3.5	*CHD7*	c.6217C>T, p.Q2073X
CHARGE	MS1014	m	0.01	*CHD7*	c.5050+1G>A
CHARGE	MS1015	m	0.7	*CHD7*	c.5405-17G>A, abnormal splicing confirmed by RNA-Seq
CHARGE	MS1016	m	17.5	*CHD7*	c.7047C>A, p.Y2349X
CHARGE	MS1019	f	0.01	*CHD7*	c.7252C>T, p.R2418X
CHARGE	MS1021	f	8.7	*CHD7*	c.5627C>G, p.S1876X
CHARGE	MS1022	m	0.01	*CHD7*	c.1924_1931del8
CHARGE	MS1023	m	0.5	*CHD7*	c.7879C>T, p.R2627X
CHARGE	MS1024	m	0.001	*CHD7*	c.6044delT
CHARGE	MS1027	f	0.3	*CHD7*	c.3339delG
CHARGE	MS1029	m	0.1	*CHD7*	c.4402delG

CHARGE	MS1030	m	0.2	*CHD7*	c.5074G>T, p.G1692X
CHARGE	MS1032	m	1.5	*CHD7*	c.3548 delA
CHARGE	MS1035	m	0.2	*CHD7*	c.6243C>A, p.C2081X).
CHARGE	MS1036	f	0.9	*CHD7*	c.5768dupG
CHARGE	MS1037	f	1	*CHD7*	c.7282C>T, p.R2428X
CHARGE	MS1038	m	2.1	*CHD7*	c.1152delA
CHARGE	MS1039	f	0.1	*CHD7*	c.2614-2_2614-1dupAG
CHARGE	MS1044	m	0.1	*CHD7*	c.5574delA
CHARGE	MS1045	m	0.2	*CHD7*	c.4862G>A, p.W1621X
CHARGE	MS1046	f	0.1	*CHD7*	c.3359delG
CHARGE	MS1049	m	10.1	*CHD7*	7219delA
Coffin-Siris Syndrome	MS0674	m		*ARID1B*	
Coffin-Siris Syndrome	MS0676	f		*ARID1B*	
Coffin-Siris Syndrome	MS0678	f		*ARID1B*	
Coffin-Siris Syndrome	MS0679	f		*SMARCB1*	
Coffin-Siris Syndrome	MS0680	m		*ARID1B*	
Coffin-Siris Syndrome	MS0681	f		*SMARCB1*	
Coffin-Siris Syndrome	MS0683	m		*SMARCB1*	
Coffin-Siris Syndrome	MS0684	m		*ARID1B*	
Coffin-Siris Syndrome	MS0685	f		*ARID1B*	
ADCA-DN	MS0494	f		*DNMT1*	DNMT1 - NM_001130823.1:c.1709C>T [p.Ala570Val]
ADCA-DN	MS0495	f		*DNMT1*	DNMT1 - NM_001130823.1:c.1709C>T [p.Ala570Val]
ADCA-DN	MS0496	m		*DNMT1*	DNMT1 - NM_001130823.1:c.1709C>T [p.Ala570Val]
ADCA-DN	MS0497	m		*DNMT1*	DNMT1 - NM_001130823.1:c.1709C>T [p.Ala570Val]
ADCA-DN	MS0498	m		*DNMT1*	DNMT1 - NM_001130823.1:c.1709C>T [p.Ala570Val]
Kabuki	KMT2D-1	f	15	*KMT2D*	c.15061C>T (p.Arg5021*)
Kabuki	KMT2D-10	m	6	*KMT2D*	c.4135_4136delA (p.Met1379Valfs*52)
Kabuki	KMT2D-11	m	16	*KMT2D*	c.11710C>T (p.Gln3904*)
Kabuki	KMT2D-2	m	1	*KMT2D*	c.16318delG (p.Glu5440Argfs*16)

Kabuki	KMT2D-22	f		*KMT2D*	c.11158C>T (p.Gln3720*)
Kabuki	KMT2D-25	f		*KMT2D*	c.16521+1G>T
Kabuki	KMT2D-26	f		*KMT2D*	c.1940dupC (p.Pro648Thrfs*2)
Kabuki	KMT2D-29	m		*KMT2D*	c.15088C>T (p.Arg5030Cys)
Kabuki	KMT2D-3	m	18	*KMT2D*	c15030dupA (p.Glu5011Argfs*13)
Kabuki	KMT2D-30	m		*KMT2D*	c.5135delA (p.Lys1712Argfs*10)
Kabuki	KMT2D-32	m		*KMT2D*	c.16052G>A (p.Arg5351Gln)
Kabuki	KMT2D-33	f		*KMT2D*	c.11203C>T (p.Gln3735*)
Kabuki	KMT2D-36	f		*KMT2D*	c.15536G>A (p.Arg5179His)
Kabuki	KMT2D-4	f	16	*KMT2D*	c.8172_8173delC (p.Phe2724Glnfs*5)
Kabuki	KMT2D-5	m	15	*KMT2D*	c.6595delT (p.Tyr2199Ilefs*65)
Kabuki	KMT2D-6	m	11	*KMT2D*	c.14055-14056delCA (p.His4685Glnfs*4)
Kabuki	KMT2D-7	m	14	*KMT2D*	c.6295C>T (p.Arg2099*)
Kabuki	KMT2D-8	m	20	*KMT2D*	c.4135_4136delA (p.Met1379Valfs*52)
Kabuki	KMT2D-9	m	18	*KMT2D*	c.12592C>T (p.Arg4198*)
Kabuki	MS0608	m	7	*KMT2D*	MLL2 c.1801_1822dup 22
Kabuki	MS0711	f	6.7	*KMT2D*	c.13059delG
Kabuki	MS0713	m	1.9	*KMT2D*	c.839+1delG
Kabuki	MS0719	f	3.9	*KMT2D*	c.15844C>T (p.R5282X)
Kabuki	MS0720	m	21.6	*KMT2D*	c.16294C>T (p.R5432W)
Kabuki	MS0721	f	0	*KMT2D*	c.8488C>T (p.R2830X)
Kabuki	MS0738	f	3.8	*KMT2D*	c.4168dupG
Kabuki	MS0740	m	4.3	*KMT2D*	c.15289C>T (p.R5097X)
Kabuki	MS0744	m	2.6	*KMT2D*	c.4419-2A>G
Kabuki	MS0774	f	16.6	*KMT2D*	c.548delC
Kabuki	MS0776	f	3.3	*KMT2D*	c.7411C>T (p.R2471X)
Kabuki	MS0778	f	24.1	*KMT2D*	c.1966dupC
Kabuki	MS0782	f	9.5	*KMT2D*	c.6200delA
Kabuki	MS0783	f	9.3	*KMT2D*	c.7933C>T (p.R2645X)
Kabuki	MS0784	f	5.8	*KMT2D*	c.13450C>T (p.R4484X)

Kabuki	MS0786	f	19.1	*KMT2D*	c.16048A>T (p.K5350X)
Kabuki	MS0787	m	7.1	*KMT2D*	c.10201C>T (p.Q3401X)
Kabuki	MS0788	m	3.4	*KMT2D*	c.16360C>T (p.R5454X)
Kabuki	MS0790	m	3.1	*KMT2D*	c.8692C>T (p.Q2898X)
Kabuki	MS0792	f	4.1	*KMT2D*	c.14878C>T (p.R4960X)
Kabuki	MS0793	f	23.1	*KMT2D*	c.6265A>T (p.K2089X)
Kabuki	MS0794	f	6.9	*KMT2D*	c.10740+1G>A
Kabuki	MS0795	m	2.2	*KMT2D*	c.13652T>A (p.L4551X)
Kabuki	MS0796	f	1	*KMT2D*	c.11596C>T (p.Q3866X)
Kabuki	MS0842	f	4	*KMT2D*	c. 4135_4136delAT
Claes-Jensen	MS0690	m	30	*KDM5C*	c.1510G>A; p.V504M
Claes-Jensen	MS0691	m	6	*KDM5C*	c.1439C>T; p.P480L
Claes-Jensen	MS0692	m	26	*KDM5C*	c.4439_4440delAG; p.fsR1481fxX9
Claes-Jensen	MS0693	m	42	*KDM5C*	IVS11+5G>A; p.E468GfsX2
Claes-Jensen	MS0695	m	8	*KDM5C*	c.1510G>A; p.V504M
Claes-Jensen	MS0697	m	2	*KDM5C*	c.1439C>T; p.P480L
Claes-Jensen	MS0723	m	37	*KDM5C*	c.229G>A; p.A77T
Claes-Jensen	MS0726	m	28	*KDM5C*	c.4439_4440delAG; p.fsR1481fxX9
Claes-Jensen	MS0727	m	13	*KDM5C*	c.229G>A; p.A77T
Claes-Jensen	MS0728	m	26	*KDM5C*	c.1510G>A; p.V504M
Rett	MS0516	f	1	*MECP2*	R106W
Rett	MS0517	f	25	*MECP2*	R168X
Rett	MS0518	f	34	*MECP2*	P302R
Rett	MS0519	f	2	*MECP2*	c1085del1113
Rett	MS0520	f	1	*MECP2*	T158M
Rett	MS0521	f	3	*MECP2*	917_939del123
Rett	MS0522	f	3	*MECP2*	deletion in exon 4
Rett	MS0523	f	1	*MECP2*	T158M
Rett	MS0524	f	4	*MECP2*	P225R
Rett	MS0525	f	6	*MECP2*	1157_1197 del 41c (c terminal deletion)

Rett	MS0526	f	1.5	*MECP2*	R255X
Rett	MS0527	m	0.7	*MECP2*	507 insAG
Rett	MS0555	f	6	*MECP2*	deletion in exon 3 and 4
Rett	MS0557	f	29	*MECP2*	R106W
Rett	MS0559	f	3	*MECP2*	T158M
Rett	MS0561	f	11	*MECP2*	R255X
Rett	MS0563	f	4	*MECP2*	partial deletion of exon 4
Coffin-Lowry	MS0513	f	6	*RSK2*	c.1520insA; p. fs507X
Coffin-Lowry	MS0514	m		*RSK2*	c.1894C>T; p. R632X
Coffin-Lowry	MS0515	m	11.5	*RSK2*	c.371G>S; p.fs372X
Coffin-Lowry	MS0528	m	11.5	*RSK2*	c.2065C>T; p.Q688X
Coffin-Lowry	MS0530	m	4	*RSK2*	c.2186G>A; p. R729Q
Coffin-Lowry	MS0532	m	1.3	*RSK2*	c.340C>T; p.W114R
Coffin-Lowry	MS0533	m	13	*RSK2*	c.727C>T; p. X243
Coffin-Lowry	MS0534	m	22.8	*RSK2*	IVS14+1G>A
Coffin-Lowry	MS0535	m	12	*RSK2*	c.62bp deletion spanning 35bp at 3'end of intron 7 and 29 bp at 5' of exon 8
Coffin-Lowry	MS0536	m	2	*RSK2*	c.386insCTTT; p.fs 130 X141
Coffin-Lowry	MS0538	m	8	*RSK2*	c.1155delT; p. fs385-X40 codons downstream
Sotos	11D/6637	m	10	*NSD1*	ex15-19 del
Sotos	11D/6718	f	10	*NSD1*	c.1716delC (p.Cys573Valfs*26)
Sotos	11D_0326	f	9	*NSD1*	chr5:175,366,008-177,470,488 (hg19)
Sotos	11D_0328	f	7	*NSD1*	chr5:175,764,262-177,059,256 (hg19)
Sotos	A1208	f	3.5	*NSD1*	c.6454C>T (p.Arg2152*)
Sotos	DL117330	f	13.2	*NSD1*	c.5445C>G (p.Tyr1815*)
Sotos	DL122057	m	3	*NSD1*	c.4843delT (p.Tyr1615Thrfs*27)
Sotos	DL151889	m	2.2	*NSD1*	microdeletion of distal 5q35.2
Sotos	DL159249	f	12	*NSD1*	c.6349C>T (p.Arg2117*)
Sotos	DL168744	m	2.2	*NSD1*	c.1492C>T (p.Arg498*)
Sotos	DL179067	m	18	*NSD1*	c.6454C>T (p.Arg2152*)
Sotos	DL38402	m	19.7	*NSD1*	c.1583delA (p.Lys528Argfs*8)

Sotos	DL50448	m	8	*NSD1*	c.2014-2018delACAGA (p.Thr672Glufs*9)
Sotos	DL50450	m	41	*NSD1*	c.2014-2018delACAGA (p.Thr672Glufs*9)
Sotos	DL50452	f	2	*NSD1*	c.2014-2018delACAGA (p.Thr672Glufs*9)
Sotos	DL76010	f	1.6	*NSD1*	c.1810C>T (p.Arg604*)
Sotos	DL87406	m	0.5	*NSD1*	5q35.2-35.3 (RP11-67P18 to RP11-423H2), FISH BAC RP11-99N22 deleted
Sotos	DL89813	m	10.6	*NSD1*	c.1801A>T (p.Lys601*)
Sotos	DL94609	m	20	*NSD1*	4977_4978insG (p.Arg1660Alafs*13)
Sotos	HK-10067	m		*NSD1*	c.6013C>T (p.Arg2005*)
Sotos	HK-11302	f		*NSD1*	c.5397_c.5398insT (p.Gly1800Trpfs*3)
Sotos	HK-11807	f		*NSD1*	c.5887A>T (p.Lys1963*)
Sotos	HK-13565	f		*NSD1*	c.2607_c.2608insA (p.Arg870Lysfs*5)
Sotos	HK-15164	f		*NSD1*	WGD-Whole gene deletion
Sotos	HK-1540	m		*NSD1*	WGD-Whole gene deletion
Sotos	HK-15438	f		*NSD1*	WGD-Whole gene deletion
Sotos	HK-2129	f		*NSD1*	c.4417C>T (p.Arg1473*)
Sotos	HK-4470	f		*NSD1*	c.1754C>G (p.Ser585*)
Sotos	HK-4913	f		*NSD1*	c.2362C>T (p.Arg788*)
Sotos	HK-5107	f		*NSD1*	WGD-Whole gene deletion
Sotos	HK-5439	f		*NSD1*	c.3728delG (p.Ser1243Thrfs*14)
Sotos	HK-5530	m		*NSD1*	c.2334_c.2335insA (p.His779Thrfs*30)
Sotos	HK-5533	m		*NSD1*	c.2314delG (p.Ala772Glnfs*19)
Sotos	HK-5589	f		*NSD1*	c.1310C>G (p.Ser437*)
Sotos	HK-5977	f		*NSD1*	c.6349C>T (p.Arg2117*)
Sotos	HK-6779	m		*NSD1*	c.3067C>T (p.Arg1023*)
Sotos	HK-6966	f		*NSD1*	c.5296C>A (R1766X p.= (p.Arg1766Arg))
Sotos	HK-7783	m		*NSD1*	c.4217delGAAA (p.Arg1406Asnfs*12)
Floating Harbor	MS0433	m		*SRCAP*	
Floating Harbor	MS0434	f		*SRCAP*	
Floating Harbor	MS0435	m		*SRCAP*	
Floating Harbor	MS0437	f		*SRCAP*	

Floating Harbor	MS0439	f		*SRCAP*	
Floating Harbor	MS0440	f		*SRCAP*	
Floating Harbor	MS0441	f		*SRCAP*	
Floating Harbor	MS0442	f		*SRCAP*	
Floating Harbor	MS0443	m		*SRCAP*	
Floating Harbor	MS0444	m		*SRCAP*	
Floating Harbor	MS0445	m		*SRCAP*	
Floating Harbor	MS0446	f		*SRCAP*	
Floating Harbor	MS0447	f		*SRCAP*	
Floating Harbor	MS0449	f		*SRCAP*	
Floating Harbor	MS0450	f		*SRCAP*	
Floating Harbor	MS0451	m		*SRCAP*	
Floating Harbor	MS0452	f		*SRCAP*	
Saethre-Chotzen Syndrome	MS0611	f	0.003	*TWIST*	c.385_405dup21
Saethre-Chotzen Syndrome	MS0612	m	5.7	*TWIST*	c.261del13
Saethre-Chotzen Syndrome	MS0613	f	0.9	*TWIST*	c.156delC
Saethre-Chotzen Syndrome	MS0614	m	38	*TWIST*	A G>T transversion at nucleotide 376 was detected (E126X)
Saethre-Chotzen Syndrome	MS0615	m	30	*TWIST*	A 21 basepair insertion beginning at nucleotide 406 was detected.
Saethre-Chotzen Syndrome	MS0617	f	33.5	*TWIST*	The deletion of a cytosine at nucleotide 156 was detected.
Saethre-Chotzen Syndrome	MS0618	m	17.7	*TWIST*	Insertion of 21 base pairs beginning at nucleotide 418
Saethre-Chotzen Syndrome	MS0624	f	20.7	*TWIST*	c.211C>T, p.Q71X
Saethre-Chotzen Syndrome	MS0625	m	0.7	*TWIST*	c.325C>T, p.Q109X
Saethre-Chotzen Syndrome	MS0626	m	0.1	*TWIST*	c.396_416dup21
Saethre-Chotzen Syndrome	MS0632	f	0.01	*TWIST*	c.193G>T, p.E65X
Saethre-Chotzen Syndrome	MS0633	f	23.3	*TWIST*	c.472T>C, p.F158L
Saethre-Chotzen Syndrome	MS0638	f	0.35	*TWIST*	entire TWIST gene deleted
Saethre-Chotzen Syndrome	MS0648	f	0.003	*TWIST*	MLPA analysis detected a full gene deletion involving the TWIST gene.
Saethre-Chotzen Syndrome	MS0653	f	0.7	*TWIST*	A G>T transition at nucleotide 160 (stop codon for a glycine at amino acid 54)
Saethre-Chotzen Syndrome	MS0654	f	20.5	*TWIST*	c.397_417dup21
Saethre-Chotzen Syndrome	MS0656	m	0.6	*TWIST*	c.120_145del26

Saethre-Chotzen Syndrome	MS0657	f	23.5	*TWIST*	c.149delC
Saethre-Chotzen Syndrome	MS0660	f	12.3	*TWIST*	A 21 base pair deletion beginning at nucleotide 394 was detected.
Saethre-Chotzen Syndrome	MS0665	f	21.5	*TWIST*	c.352C>G, p.R118G
Saethre-Chotzen Syndrome	MS0668	f	0.8	*TWIST*	c.376G>T, p.Gln126X
Saethre-Chotzen Syndrome	MS0669	f	28.7	*TWIST*	c.490C>T, p.Q164X
Saethre-Chotzen Syndrome	MS0688	m	0.02	*TWIST*	c.149delC
Saethre-Chotzen Syndrome	MS0689	f	0.1	*TWIST*	c.149delC
Saethre-Chotzen Syndrome	MS0844	f	23.5	*TWIST*	full gene deletion of TWIST
Weaver	A117	f		*EZH2*	c.394C>T (p.Pro132Ser)
Weaver	A123	m		*EZH2*	c.457_459del (p.Tyr153del)
Weaver	A134	f		*EZH2*	c.2080C>T (p.His694Tyr)
Weaver	A1359	m		*EZH2*	c.2050C>T (p.Arg684Cys)
Weaver	A212	f		*EZH2*	c.398A>G (p.Tyr133Cys)
Weaver	DL129940	m		*EZH2*	c.1876G>A (p.Val626Met)
Weaver	F0123	m		*EZH2*	c.553G>C (p.Asp185His)
GTPTS	MS0673	f		*KAT6B*	
GTPTS	MS0675	m		*KAT6B*	
GTPTS	MS0677	m		*KAT6B*	
SBBYSS	MS0682	f		*KAT6B*	

Age and mutation information from some of the samples have been masked to protect the identities. Reference sequence for all mutations is hg19, except for NSD1 mutations (hg18, unless specified otherwise [11D-0326 and 11D-0328])

Table S2 – Probes shared by the epi-signatures of more than two conditions

Probe	Chr	Position	Genes	Enhancer	CpG island	ADCA-DN	Floating Harbor	ATRX	Kabuki	CHARGE	Claes-Jensen	Sotos	GTPTS	SBBYSS
cg20768358	chr1	3574999	*TP73*		OpenSea			Yes	Yes			Yes		
cg04015962	chr1	10949192			OpenSea				Yes		Yes	Yes		
cg06790069	chr1	40203513	*PPIE*		N_Shore		Yes			Yes				Yes
cg02283735	chr1	44011587	*PTPRF*	TRUE	OpenSea				Yes		Yes	Yes		
cg06766016	chr1	45278971	*BTBD19*		Island		Yes				Yes	Yes		
cg19530728	chr1	45279132	*BTBD19*		Island		Yes				Yes	Yes		
cg04605287	chr1	54953486			N_Shore				Yes		Yes	Yes		
cg08128444	chr1	68966466			S_Shelf			Yes		Yes		Yes		
cg20457732	chr1	71172486		TRUE	OpenSea		Yes				Yes	Yes		
cg03847896	chr1	112154295		TRUE	OpenSea		Yes	Yes		Yes				
cg25066665	chr1	150335507	*RPRD2*		N_Shore		Yes				Yes	Yes		
cg15448220	chr1	150897856	*SETDB1*		N_Shore		Yes				Yes	Yes		
cg10589385	chr1	150898437	*SETDB1*		N_Shore		Yes				Yes	Yes		
cg13069100	chr1	155035008	*ADAM15;EFNA4*		Island		Yes			Yes				Yes
cg09938479	chr1	178455912			Island		Yes		Yes		Yes	Yes		
cg26004771	chr1	178456093			Island		Yes				Yes	Yes		
cg22790377	chr1	228291668	*C1orf35*		S_Shore		Yes	Yes	Yes					
cg20462561	chr1	228890820			Island		Yes		Yes			Yes		
cg03982998	chr1	238024671	*LOC100130331*		OpenSea			Yes		Yes		Yes		
cg11166453	chr1	247681781			Island		Yes				Yes	Yes		Yes
cg06362358	chr1	248855545			S_Shore		Yes	Yes	Yes					
cg03701930	chr10	1981436			OpenSea				Yes	Yes		Yes		
cg08584759	chr10	11912126	*C10orf47*		Island				Yes	Yes		Yes		
cg20137746	chr10	27235598			OpenSea		Yes	Yes				Yes		
cg01001533	chr10	44184868			N_Shore				Yes			Yes	Yes	
cg24530234	chr10	44224014			Island		Yes	Yes				Yes		
cg12845268	chr10	63657363		TRUE	OpenSea				Yes		Yes	Yes		

cg09411922	chr10	75121857			S_Shelf			Yes		Yes		Yes		
cg12667595	chr10	81743362			Island		Yes		Yes			Yes		
cg16844333	chr10	89167457			OpenSea	Yes	Yes					Yes		
cg23615741	chr10	101297642		TRUE	Island		Yes		Yes				Yes	Yes
cg27245056	chr10	122356491			OpenSea		Yes	Yes		Yes				
cg07040013	chr10	132099553			OpenSea	Yes			Yes			Yes		
cg22079043	chr11	1785631	*CTSD;HCCA2*		S_Shore		Yes				Yes	Yes		
cg13924715	chr11	10750890			OpenSea			Yes	Yes			Yes		
cg05488217	chr11	58730897		TRUE	N_Shore		Yes			Yes			Yes	
cg17379860	chr11	61159602	*TMEM216*		N_Shore			Yes			Yes	Yes		
cg04638150	chr11	62273418	*AHNAK*	TRUE	OpenSea			Yes	Yes			Yes		
cg01178624	chr11	65360327	*KCNK7*		Island		Yes		Yes					Yes
cg22203628	chr11	69241075			OpenSea		Yes			Yes		Yes		
cg04006327	chr11	72533487	*ATG16L2*	TRUE	Island		Yes			Yes				Yes
cg03274739	chr11	98437942			OpenSea		Yes			Yes		Yes		
cg06137123	chr11	129444480			OpenSea			Yes	Yes	Yes		Yes		
cg05237469	chr11	133446415			OpenSea			Yes	Yes		Yes			
cg10635895	chr12	2339439	*CACNA1C*		Island		Yes				Yes	Yes		
cg01837362	chr12	34492938			N_Shore			Yes	Yes			Yes		
cg16615454	chr12	50475167	*ACCN2*		Island			Yes				Yes	Yes	
cg03146625	chr12	54448729	*HOXC4*		S_Shore				Yes	Yes		Yes		
cg17970299	chr12	54772804	*ZNF385A*	TRUE	Island		Yes	Yes	Yes					
cg08425810	chr12	58132558	*AGAP2*		Island				Yes	Yes		Yes		
cg00047843	chr12	83321412	*TMTC2*	TRUE	OpenSea			Yes		Yes		Yes		
cg10082647	chr12	107348855	*C12orf23*		N_Shore						Yes	Yes	Yes	
cg11367159	chr12	110044531		TRUE	OpenSea		Yes				Yes	Yes		
cg10504392	chr12	110044639		TRUE	OpenSea		Yes				Yes	Yes		
cg13810723	chr12	119592343	*SRRM4*		S_Shore	Yes		Yes				Yes		
cg08083251	chr13	23309930			OpenSea		Yes		Yes			Yes		
cg13614409	chr13	31506752	*C13orf26*		OpenSea		Yes		Yes			Yes		

cg26168643	chr13	88328009	*SLITRK5*		N_Shore			Yes	Yes	Yes				
cg24626752	chr13	88328274	*SLITRK5*		N_Shore				Yes	Yes				Yes
cg05757365	chr13	88328471	*SLITRK5*		N_Shore				Yes	Yes				Yes
cg08602346	chr13	113242997	*TUBGCP3*		S_Shore						Yes	Yes	Yes	
cg14009504	chr14	71555837	*PCNX*	TRUE	OpenSea			Yes		Yes		Yes		
cg09214243	chr15	29968124			S_Shore				Yes		Yes	Yes		
cg12414681	chr15	29968195			S_Shore				Yes		Yes	Yes		
cg09456065	chr15	56759655			S_Shelf			Yes		Yes		Yes		
cg02703787	chr15	93256789			N_Shore		Yes	Yes					Yes	
cg27009812	chr16	3062597	*CLDN9*	TRUE	N_Shore					Yes	Yes	Yes		
cg04715503	chr16	3062795	*CLDN9*	TRUE	N_Shore					Yes	Yes	Yes		
cg00524708	chr16	31159558	*PRSS36*		N_Shore			Yes	Yes			Yes		
cg10424751	chr16	32289757			Island			Yes	Yes			Yes		
cg00130223	chr16	33070551			Island				Yes	Yes		Yes		
cg03203197	chr16	34587120			Island				Yes	Yes		Yes		
cg01517680	chr16	49499006		TRUE	Island		Yes	Yes				Yes		
cg12499872	chr16	58019893	*TEPP*		Island	Yes		Yes				Yes		
cg09482050	chr16	67686832	*RLTPR*		N_Shore		Yes	Yes				Yes		
cg06911613	chr16	85846184			S_Shore				Yes			Yes		Yes
cg04387835	chr17	4649076	*ZMYND15*		OpenSea			Yes	Yes			Yes		
cg17900689	chr17	4649262	*ZMYND15*		OpenSea			Yes	Yes			Yes		
cg03613822	chr17	7115140	*DLG4*		N_Shelf		Yes	Yes	Yes					
cg19786602	chr17	7966326			OpenSea			Yes	Yes			Yes		
cg06663305	chr17	8095813			S_Shelf		Yes	Yes	Yes					
cg16006004	chr17	41278241	*BRCA1;NBR2*		Island		Yes		Yes					Yes
cg07075026	chr17	47091521	*IGF2BP1*		Island		Yes			Yes		Yes		
cg10950924	chr17	47092072	*IGF2BP1*	TRUE	Island		Yes			Yes		Yes		
cg18040354	chr17	53800484	*TMEM100*		OpenSea		Yes			Yes	Yes			
cg12131208	chr17	58499700	*C17orf64*	TRUE	S_Shore				Yes		Yes	Yes		
cg21122199	chr17	58499720	*C17orf64*	TRUE	S_Shore	Yes					Yes	Yes		

cg08122232	chr17	74100132	*EXOC7*		S_Shore		Yes				Yes	Yes	Yes	
cg12082025	chr19	1064218	*ABCA7*		Island		Yes	Yes				Yes		
cg14576825	chr19	1676136			Island	Yes							Yes	Yes
cg18795569	chr19	3480508	*C19orf77*		N_Shore		Yes	Yes				Yes		
cg02085507	chr19	6739192	*TRIP10*		N_Shore		Yes		Yes			Yes		
cg22987448	chr19	8591364	*MYO1F*		Island		Yes		Yes	Yes				
cg08283130	chr19	8591776	*MYO1F*		Island		Yes		Yes	Yes				
cg14266237	chr19	12995624	*KLF1*		N_Shore		Yes				Yes			Yes
cg19492423	chr19	18700933	*C19orf60*		S_Shore		Yes	Yes	Yes		Yes			Yes
cg21977234	chr19	18701320	*C19orf60*		S_Shore		Yes	Yes			Yes			Yes
cg26931208	chr19	49244592	*IZUMO1*		S_Shore		Yes	Yes				Yes		
cg26750002	chr19	49244674	*IZUMO1;RASIP1*		S_Shore		Yes	Yes				Yes		
cg17434634	chr19	49522954		TRUE	Island		Yes						Yes	Yes
cg05934682	chr19	50394298	*IL4I1*	TRUE	S_Shore			Yes	Yes			Yes		
cg11383165	chr19	56002283	*SSC5D*		Island				Yes			Yes		Yes
cg16151651	chr2	1609119			OpenSea		Yes				Yes	Yes		
cg18396357	chr2	50201413	*NRXN1*		OpenSea		Yes	Yes		Yes				
cg09692695	chr2	74668286	*RTKN*		Island			Yes			Yes	Yes		
cg25334934	chr2	121269348		TRUE	OpenSea				Yes		Yes	Yes		
cg18478731	chr2	129659682			Island		Yes		Yes			Yes		
cg11942181	chr2	129659946			Island		Yes		Yes			Yes		
cg09015973	chr2	131673771	*ARHGEF4*		Island				Yes			Yes		Yes
cg03727500	chr2	232348334		TRUE	Island				Yes		Yes	Yes		
cg11559198	chr2	232348794			Island				Yes		Yes	Yes		
cg05857996	chr20	2675418	*EBF4*	TRUE	S_Shore				Yes		Yes	Yes		
cg24263062	chr20	2730191	*EBF4*		Island				Yes		Yes	Yes		
cg12176783	chr20	62694000	*TCEA2*		Island		Yes		Yes					Yes
cg05437132	chr21	10990903	*TPTE*		Island			Yes		Yes		Yes		
cg04767174	chr21	15069190			Island			Yes	Yes			Yes		
cg08597025	chr21	15077096			Island			Yes		Yes		Yes		

cg21916358	chr21	15077562			Island			Yes	Yes	Yes		Yes		
cg10044179	chr21	15352983	*C21orf81*		S_Shore			Yes	Yes	Yes		Yes		
cg23720384	chr21	23747055			OpenSea			Yes	Yes				Yes	
cg14172108	chr21	34405553			N_Shore				Yes		Yes	Yes		
cg00274965	chr21	34405681			Island				Yes		Yes	Yes		
cg03389890	chr22	16868045			Island				Yes			Yes	Yes	
cg16054907	chr22	50706065	*MAPK11*		Island		Yes	Yes		Yes				
cg07225641	chr3	46618325	*LRRC2;TDGF1*		Island	Yes						Yes	Yes	
cg08234664	chr3	49170668	*LAMB2*		OpenSea		Yes		Yes		Yes	Yes		
cg14099457	chr3	49170794	*LAMB2*		OpenSea				Yes		Yes	Yes		
cg10313047	chr3	52351355	*DNAH1*		OpenSea		Yes		Yes			Yes		
cg00487142	chr3	109056357	*DPPA4*		OpenSea		Yes	Yes				Yes		
cg14836960	chr3	109056618	*DPPA4*		OpenSea		Yes	Yes				Yes		
cg21750709	chr3	109056897	*DPPA4*		OpenSea		Yes	Yes				Yes		
cg01571001	chr3	128215433			Island		Yes		Yes					Yes
cg01757168	chr3	128215565			Island		Yes		Yes	Yes				Yes
cg09406615	chr3	165555103	*BCHE*	TRUE	OpenSea		Yes					Yes	Yes	
cg10663765	chr3	194014592			Island				Yes		Yes	Yes		
cg23855319	chr4	2062392	*NAT8L*		Island		Yes				Yes	Yes		
cg16091553	chr4	46126245	*GABRG1*	TRUE	OpenSea		Yes			Yes	Yes			
cg19676181	chr4	56660398			S_Shore		Yes					Yes	Yes	
cg06013872	chr4	72203823	*SLC4A4*		OpenSea		Yes	Yes		Yes				
cg03670393	chr4	81128690		TRUE	Island	Yes	Yes					Yes		
cg04511182	chr4	103269584			S_Shelf			Yes		Yes		Yes		
cg13023646	chr4	165898835	*TRIM61*		OpenSea		Yes					Yes	Yes	
cg11685843	chr4	176349171			OpenSea		Yes	Yes	Yes	Yes				
cg09166973	chr5	23507573	*PRDM9*		OpenSea			Yes			Yes	Yes		
cg07011961	chr5	23507594	*PRDM9*		OpenSea		Yes	Yes		Yes	Yes	Yes		
cg25472530	chr5	23507617	*PRDM9*		OpenSea		Yes	Yes			Yes			
cg22079902	chr5	23507644	*PRDM9*		OpenSea			Yes			Yes	Yes		

cg01667892	chr5	23507656	*PRDM9*		OpenSea			Yes			Yes	Yes		
cg23688479	chr5	42756851	*CCDC152*		OpenSea		Yes	Yes				Yes		
cg16026035	chr5	70742893			OpenSea			Yes			Yes	Yes		
cg06501366	chr5	78365687	*BHMT2*	TRUE	Island		Yes	Yes				Yes		
cg08328513	chr5	78365691	*BHMT2;DMGDH*	TRUE	Island		Yes	Yes				Yes		
cg04117076	chr5	88991074		TRUE	OpenSea			Yes		Yes		Yes		
cg14575983	chr5	99922152	*FAM174A*		OpenSea			Yes		Yes		Yes		
cg18665594	chr5	101119420			OpenSea		Yes		Yes			Yes		
cg19526166	chr5	110062729	*TMEM232*		OpenSea		Yes	Yes			Yes	Yes		
cg26817546	chr5	135538704		TRUE	OpenSea		Yes				Yes	Yes		
cg08951638	chr5	141706213		TRUE	S_Shore	Yes	Yes					Yes		
cg19731612	chr5	176559334	*NSD1*	TRUE	Island		Yes				Yes	Yes		
cg17493885	chr5	176559558	*NSD1*		Island		Yes				Yes	Yes		
cg18121224	chr5	176559563	*NSD1*		Island		Yes				Yes	Yes		
cg26092675	chr6	26225258	*HIST1H3E*		N_Shore		Yes	Yes			Yes			Yes
cg13836098	chr6	26225268	*HIST1H3E*		N_Shore		Yes	Yes			Yes			Yes
cg22040809	chr6	26522578	*HCG11*		Island	Yes					Yes	Yes		
cg17762073	chr6	34024220	*GRM4*	TRUE	Island			Yes			Yes	Yes		
cg20646500	chr6	42536105	*UBR2*		S_Shelf		Yes	Yes		Yes				
cg17097119	chr6	49917872	*DEFB133*		OpenSea			Yes		Yes		Yes		
cg00389785	chr6	55035173			N_Shelf		Yes	Yes		Yes		Yes		
cg18090145	chr6	67741714			OpenSea			Yes	Yes	Yes				
cg14079463	chr6	127796989	*C6orf174*		Island		Yes	Yes			Yes			
cg04580344	chr6	127797022	*C6orf174*		Island	Yes		Yes			Yes			
cg08883853	chr7	1026168	*CYP2W1*		Island					Yes		Yes		Yes
cg12995372	chr7	11066024	*PHF14*		OpenSea			Yes		Yes		Yes		
cg24231558	chr7	16626326			Island			Yes		Yes		Yes		
cg18081818	chr7	23246105			Island						Yes	Yes	Yes	
cg05407003	chr7	23246146			Island						Yes	Yes	Yes	
cg11015251	chr7	27170554	*HOXA4*	TRUE	Island				Yes	Yes			Yes	

cg20974609	chr7	27181671	*HOXA5*		N_Shore				Yes	Yes			Yes	
cg11724970	chr7	27182493	*HOXA5*		N_Shore				Yes	Yes			Yes	
cg05076221	chr7	27182637	*HOXA5*		Island				Yes	Yes			Yes	
cg23936031	chr7	27183133	*HOXA5*		Island				Yes	Yes			Yes	
cg04863892	chr7	27183375	*HOXA5*		Island				Yes	Yes			Yes	
cg19759481	chr7	27183401	*HOXA5*		Island				Yes	Yes			Yes	
cg02916332	chr7	27183591	*HOXA5*		Island				Yes	Yes			Yes	
cg17569124	chr7	27183643	*HOXA5*		Island				Yes	Yes			Yes	
cg02005600	chr7	27183686	*HOXA5*		Island				Yes	Yes			Yes	
cg25307665	chr7	27183694	*HOXA5*		Island				Yes	Yes			Yes	
cg14014955	chr7	27183701	*HOXA5*		Island				Yes	Yes			Yes	
cg02646423	chr7	27183794	*HOXA5*		Island				Yes	Yes			Yes	
cg20517050	chr7	27183806	*HOXA5*		Island				Yes	Yes			Yes	
cg23204968	chr7	27183816	*HOXA5*		Island				Yes	Yes			Yes	
cg20817131	chr7	27184167	*HOXA5*		Island				Yes	Yes			Yes	
cg25390165	chr7	27184188	*HOXA5*		Island				Yes	Yes			Yes	
cg14882265	chr7	27184375	*HOXA5*		Island				Yes	Yes			Yes	
cg17432857	chr7	27184438	*HOXA5*		Island	Yes			Yes	Yes			Yes	
cg00969405	chr7	27184441	*HOXA5*		Island				Yes	Yes			Yes	
cg03368099	chr7	27184521	*HOXA5*		Island				Yes	Yes			Yes	
cg27151303	chr7	27184821			Island				Yes	Yes			Yes	
cg24040595	chr7	27185512	*HOXA6*		Island				Yes	Yes			Yes	
cg23129930	chr7	27186993	*HOXA6*		Island				Yes	Yes			Yes	
cg05928186	chr7	27187102	*HOXA6*		Island				Yes	Yes			Yes	
cg02943305	chr7	31685368	*CCDC129*	TRUE	OpenSea		Yes	Yes		Yes				
cg27385590	chr7	49619841		TRUE	OpenSea		Yes					Yes	Yes	
cg26766885	chr7	96220185		TRUE	OpenSea					Yes		Yes		Yes
cg05715492	chr7	98991138	*ARPC1B*	TRUE	S_Shore		Yes		Yes					Yes
cg08234689	chr7	127910927			Island			Yes			Yes	Yes		
cg26209990	chr7	127911258			Island			Yes			Yes	Yes		

cg11688949	chr8	27115956	*STMN4*		OpenSea	Yes					Yes	Yes		
cg20893180	chr8	36957379		TRUE	OpenSea						Yes	Yes	Yes	
cg08620329	chr8	67454546		TRUE	Island		Yes	Yes				Yes		
cg12109823	chr8	67454595		TRUE	Island		Yes	Yes				Yes		
cg02173067	chr8	111214984			OpenSea		Yes					Yes	Yes	
cg12873476	chr8	142402728		TRUE	S_Shore		Yes	Yes			Yes			
cg23171972	chr8	143203477			S_Shore					Yes		Yes		Yes
cg13746854	chr9	34370894	*KIAA1161*	TRUE	Island				Yes		Yes			Yes
cg11510586	chr9	72027409		TRUE	Island		Yes		Yes			Yes		
cg14035368	chr9	139258938	*CARD9;DNLZ*		Island		Yes			Yes		Yes		

Table S3 - Top gene ontology terms found to be enriched in the genes harboring probes shared by the epi-signatures of more than two conditions

Gene Ontology	Gene Ontology Term	N	DE	P-value
GO:0007625	grooming behavior	17	3	3.80E-05
GO:2000821	regulation of grooming behavior	4	2	0.000149
GO:0097109	neuroligin family protein binding	5	2	0.000504
GO:0033130	acetylcholine receptor binding	8	2	0.0007
GO:0060484	lung-associated mesenchyme development	10	2	0.001066
GO:0060044	negative regulation of cardiac muscle cell proliferation	10	2	0.001096
GO:0018024	histone-lysine N-methyltransferase activity	43	3	0.001101
GO:0016279	protein-lysine N-methyltransferase activity	55	3	0.001883
GO:0016278	lysine N-methyltransferase activity	56	3	0.001955
GO:0042054	histone methyltransferase activity	56	3	0.00215
GO:0055022	negative regulation of cardiac muscle tissue growth	14	2	0.002192
GO:0061117	negative regulation of heart growth	14	2	0.002192
GO:0006978	DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator	16	2	0.002608
GO:0001891	phagocytic cup	19	2	0.002654
GO:0047865	dimethylglycine dehydrogenase activity	1	1	0.00278
GO:0042772	DNA damage response, signal transduction resulting in transcription	17	2	0.003039
GO:0061627	S-methylmethionine-homocysteine S-methyltransferase activity	1	1	0.003144
GO:0033477	S-methylmethionine metabolic process	1	1	0.003144
GO:0071625	vocalization behavior	15	2	0.003453
GO:0072274	metanephric glomerular basement membrane development	1	1	0.003898

N: Number of genes in the GO term; DE: number of genes that are differentially methylated

Table S4 – Pathways enriched in the genes harboring probes shared by the epi-signatures of more than two conditions

ID	Description	BgRatio	pvalue	p.adjust	qvalue	geneID	Count
R-HSA-3214841	PKMTs methylate histone lysines*	73/10281	0.000116	0.025528	0.022841	*PRDM9/SETDB1/HIST1H3E/NSD1*	4
R-HSA-6804759	Regulation of TP53 Activity through Association with Co-factors	14/10281	0.001057	0.104046	0.093094	*ZNF385A/TP73*	2
R-HSA-3000171	Non-integrin membrane-ECM interactions	64/10281	0.001419	0.104046	0.093094	*LAMB2/LAMC1/NRXN1*	3
R-HSA-6794362	Protein-protein interactions at synapses	71/10281	0.001914	0.105267	0.094186	*PTPRF/DLG4/NRXN1*	3
R-HSA-8849932	SALM protein interactions at the synapses	21/10281	0.002401	0.105641	0.094521	*PTPRF/DLG4*	2
R-HSA-3000157	Laminin interactions	33/10281	0.00588	0.184787	0.165336	*LAMB2/LAMC1*	2
R-HSA-8874081	MET activates PTK2 signaling	33/10281	0.00588	0.184787	0.165336	*LAMB2/LAMC1*	2
R-HSA-168638	NOD1/2 Signaling Pathway	36/10281	0.006969	0.191659	0.171484	*CARD9/MAPK11*	2

*Adjusted p-value<0.05

Table S5 – Probes used for development of a multiclass SVM classification model

Probe	Chr	Position	Gene	Islands_Name	Relation_to_Island
cg27619353	chr22	38071651	*LGALS1*	chr22:38073037-38073412	N_Shore
cg19386379	chr11	844284	*TSPAN4*	chr11:842293-843396	S_Shore
cg04234016	chr12	7062109	*PTPN6*		OpenSea
cg12474798	chr10	64565772	*ADO*	chr10:64564435-64565818	Island
cg02583484	chr12	54677008	*HNRNPA1*	chr12:54673322-54673550	S_Shelf
cg03010887	chr16	85600235			OpenSea
cg13759905	chr2	233741920		chr2:233740668-233741879	S_Shore
cg06768599	chr14	24785488	*LTB4R*	chr14:24785147-24785869	Island
cg04731926	chr19	35758185	*LSR*	chr19:35758201-35759670	N_Shore
cg13298859	chr7	2551996	*LFNG*		OpenSea
cg21886367	chr14	24780825	*LTB4R*	chr14:24779874-24780932	Island
cg12464638	chr11	844400	*TSPAN4*	chr11:842293-843396	S_Shore
cg07397616	chr22	38035109	*SH3BP1*	chr22:38035350-38035928	N_Shore
cg18182216	chr1	150978385	*FAM63A*	chr1:150980735-150981092	N_Shelf
cg15095327	chr3	9944512	*IL17RE*		OpenSea
cg10582687	chr8	48091424			OpenSea
cg22091609	chr22	20780508	*SCARF2*	chr22:20779386-20780560	Island
cg21721340	chr20	62682309	*SOX18*	chr20:62679424-62680883	S_Shore
cg11545720	chr19	4558377	*SEMA6B*	chr19:4558363-4558579	Island
cg17054691	chr17	79813439	*P4HB*		OpenSea
cg06247837	chr17	37820135	*TCAP*	chr17:37823693-37824989	N_Shelf
cg05551825	chr7	5735129	*RNF216*		OpenSea
cg03043406	chr1	45242356	*RPS8*	chr1:45241013-45241900	S_Shore
cg08255475	chr16	88871329	*CDT1*	chr16:88871850-88872056	N_Shore
cg20978937	chr14	105399321	*PLD4*	chr14:105398076-105400079	Island
cg01178624	chr11	65360327	*KCNK7*	chr11:65359292-65360328	Island
cg00551910	chr14	70037973	*C14orf162*	chr14:70038108-70040302	N_Shore
cg11225330	chr1	1196747	*UBE2J2*		OpenSea
cg08818610	chr6	24910720	*FAM65B*	chr6:24910626-24911285	Island
cg13710969	chr17	79393463	*BAHCC1*	chr17:79393341-79393742	Island
cg24695071	chr7	2419958	*EIF3B*	chr7:2419808-2420340	Island
cg03173827	chr20	62688717	*TCEA2*	chr20:62688551-62688878	Island
cg16823042	chr12	58119992	*AGAP2*	chr12:58119909-58121551	Island
cg06826457	chr12	12867669		chr12:12869798-12871248	N_Shelf
cg14707053	chr6	42019127	*TAF8*	chr6:42018139-42018604	S_Shore
cg20543544	chr10	81003657	*ZMIZ1*	chr10:81002109-81003687	Island
cg16565409	chr17	27048223	*RPL23A*	chr17:27046856-27047273	S_Shore
cg07211044	chr8	60032983	*TOX*	chr8:60030134-60032356	S_Shore
cg02774542	chr19	11517152	*RGL3*	chr19:11517059-11517284	Island
cg09978353	chr17	72759282	*SLC9A3R1*		OpenSea
cg10583997	chr2	32583696	*BIRC6*	chr2:32581746-32583265	S_Shore
cg26372517	chr1	36039159	*TFAP2E*	chr1:36038927-36040051	Island
cg06359132	chr10	99160096	*RRP12*	chr10:99160814-99161116	N_Shore
cg15130459	chr5	40834228	*SNORD72*	chr5:40835241-40835873	N_Shore
cg14371731	chr10	81003175	*ZMIZ1*	chr10:81002109-81003687	Island
cg27413508	chr20	30225517	*COX4I2*	chr20:30225572-30225852	N_Shore
cg14307477	chr19	51330469	*KLK15*	chr19:51330212-51330470	Island
cg22721579	chr1	36787932	*FAM176B*	chr1:36786500-36789402	Island
cg00973118	chr16	374570	*AXIN1*	chr16:374732-375328	N_Shore
cg06577850	chr7	84002356			OpenSea
cg08636573	chr7	100881842	*CLDN15*		OpenSea
cg08481464	chr10	77871618	*C10orf11*	chr10:77872083-77872305	N_Shore
cg15620905	chr1	44024150	*PTPRF*		OpenSea
cg06902698	chr1	151137676	*SCNM1*		OpenSea
cg16443812	chr22	50987294	*KLHDC7B*	chr22:50984972-50988141	Island
cg26858144	chr19	54466702	*CACNG8*	chr19:54466357-54466725	Island
cg25073708	chr9	132083538	*C9orf106*	chr9:132082871-132083582	Island

cg01832549	chr1	19774989	*CAPZB*		OpenSea
cg24869272	chr11	850296	*TSPAN4*	chr11:849087-850982	Island
cg00525011	chr16	122031	*RHBDF1*	chr16:122017-122873	Island
cg05202654	chr16	88706241	*IL17C*	chr16:88706211-88706492	Island
cg23141851	chr3	14615941		chr3:14615366-14615957	Island
cg21570209	chr19	46367987	*FOXA3*	chr19:46366202-46366605	S_Shore
cg00873601	chr12	116044025			OpenSea
cg19887750	chr18	47813174	*CXXC1*	chr18:47813434-47814546	N_Shore
cg14244402	chr9	130681102		chr9:130679120-130679612	S_Shore
cg08122386	chr6	26195954		chr6:26197070-26197537	N_Shore
cg22828581	chr9	93880502		chr9:93879725-93880618	Island
cg18385190	chr14	104940369		chr14:104940228-104940466	Island
cg05507459	chr9	91148748	*NXNL2*	chr9:91149945-91150648	N_Shore
cg19721065	chr12	113549986	*RASAL1*	chr12:113549885-113550165	Island
cg05100067	chr2	2295969	*MYT1L*		OpenSea
cg07816074	chr4	8201560	*SH3TC1*		OpenSea
cg07588439	chr2	198571554	*MARS2*	chr2:198569911-198570948	S_Shore
cg14653977	chr9	136038692	*GBGT1*	chr9:136038416-136039577	Island
cg19131476	chr11	64387923	*NRXN2*		OpenSea
cg18974224	chr14	105603153		chr14:105603310-105603532	N_Shore
cg02152233	chr1	29189789	*OPRD1*	chr1:29189307-29189952	Island
cg11499091	chr17	78735302	*RPTOR*		OpenSea
cg08498747	chr17	41747669			OpenSea
cg06951627	chr6	26196580		chr6:26197070-26197537	N_Shore
cg27539527	chr7	156735656		chr7:156735245-156735665	Island
cg08157684	chr1	27675934	*SYTL1*	chr1:27675919-27678016	Island
cg16065768	chr12	114107712			OpenSea
cg26679004	chr10	88023135	*GRID1*	chr10:88022640-88023279	Island
cg08234664	chr3	49170668	*LAMB2*		OpenSea
cg02575092	chr17	77100187	*HRNBP3*	chr17:77099995-77100203	Island
cg16875104	chr7	30635889	*GARS*	chr7:30634124-30635058	S_Shore
cg10036013	chr7	4778839	*FOXK1*		OpenSea
cg26038649	chr1	27683501	*MAP3K6*	chr1:27683277-27683590	Island
cg20956548	chr19	56618060	*ZNF787*	chr19:56615088-56615955	S_Shelf
cg05715492	chr7	98991138	*ARPC1B*	chr7:98990157-98990922	S_Shore
cg04299200	chr5	534332		chr5:535484-535814	N_Shore
cg08182193	chr14	105942243	*CRIP2*	chr14:105940325-105941900	S_Shore
cg13599248	chr6	74364700	*SLC17A5*	chr6:74363202-74364006	S_Shore
cg05164926	chr17	7255624	*KCTD11*	chr17:7254622-7255808	Island
cg01231141	chr5	178692691	*ADAMTS2*		OpenSea
cg13541527	chr6	31804078	*C6orf48*	chr6:31802598-31802823	S_Shore
cg25255429	chr17	74023633	*EVPL*		OpenSea
cg06832677	chr2	119976764			OpenSea
cg02557179	chr16	32822612		chr16:32822454-32822741	Island
cg26823666	chr1	228658397		chr1:228658937-228659152	N_Shore
cg27053299	chr13	100548780	*CLYBL*	chr13:100547633-100548911	Island
cg20698421	chr2	65217623	*SLC1A4*	chr2:65215598-65217212	S_Shore
cg10364115	chr7	2560976	*LFNG*	chr7:2558371-2559967	S_Shore
cg16017144	chr17	77100082	*HRNBP3*	chr17:77099995-77100203	Island
cg17868694	chr7	100841533	*MOGAT3*	chr7:100845050-100845726	N_Shelf
cg02380135	chr2	149866270	*KIF5C*	chr2:149866654-149866890	N_Shore
cg00497905	chr11	76903183	*MYO7A*		OpenSea
cg05620261	chr19	34991070	*WTIP*	chr19:34988748-34988951	S_Shelf
cg12996100	chr8	11470690		chr8:11471014-11471249	N_Shore
cg19392379	chr10	123909426	*TACC2*		OpenSea
cg17220161	chr1	203242601		chr1:203242519-203242818	Island
cg24247537	chr11	457278	*PTDSS2*	chr11:459088-459345	N_Shore
cg02680050	chr12	133179887		chr12:133179886-133180351	Island
cg26328687	chr11	96076794	*MAML2*		OpenSea

cg18955367	chr19	49002338	*LMTK3*	chr19:49001748-49003087	Island
cg08530537	chr16	68014374	*DPEP3*	chr16:68013967-68014545	Island
cg22523072	chr1	248300255			OpenSea
cg04359915	chr10	101825029	*CPN1*	chr10:101824961-101825186	Island
cg10555800	chr1	168356606			OpenSea
cg12845268	chr10	63657363			OpenSea
cg24996814	chr11	68417376			OpenSea
cg17740434	chr2	179388064	*MIR548N*		OpenSea
cg02548777	chr6	170338826		chr6:170338214-170339677	Island
cg19280206	chr3	196694932	*PIGZ*	chr3:196695319-196697065	N_Shore
cg23928920	chr15	92944756	*ST8SIA2*		OpenSea
cg09970481	chr19	42379744	*CD79A*		OpenSea
cg20744163	chr10	80999841	*ZMIZ1*	chr10:81002109-81003687	N_Shelf
cg16276982	chr15	29968032		chr15:29967209-29967621	S_Shore
cg08856033	chr2	121280078		chr2:121279842-121280183	Island
cg04634427	chr15	27087315		chr15:27083992-27084222	S_Shelf
cg13842556	chr10	88023243	*GRID1*	chr10:88022640-88023279	Island
cg03472000	chr10	118084320	*C10orf96*		OpenSea
cg01001533	chr10	44184868		chr10:44185039-44186132	N_Shore
cg02661079	chr20	44829722	*CDH22*	chr20:44829506-44829848	Island
cg18959207	chr10	18549686	*CACNB2*		OpenSea
cg20248611	chr12	34371861		chr12:34371627-34372304	Island
cg10800226	chr5	1850261		chr5:1851342-1851564	N_Shore
cg05348875	chr2	206628625	*NRP2*		OpenSea
cg10166889	chr5	61028374			OpenSea
cg05962028	chr12	214000	*IQSEC3*	chr12:213763-214227	Island
cg18026225	chr17	43198423	*PLCD3*	chr17:43198277-43198773	Island
cg15108641	chr10	99263320	*UBTD1*	chr10:99257940-99259520	S_Shelf
cg00284153	chr10	135171713	*C10orf125*	chr10:135170645-135171954	Island
cg10848724	chr12	123380878	*VPS37B*	chr12:123380333-123380894	Island
cg05962511	chr10	102730022		chr10:102729299-102729710	S_Shore
cg20228731	chr7	130646051	*FLJ43663*		OpenSea
cg21908488	chr17	17991580	*DRG2*	chr17:17991080-17991581	Island
cg14971320	chr10	43699070	*RASGEF1A*	chr10:43697777-43698177	S_Shore
cg02487233	chr3	107810687	*CD47*	chr3:107809375-107810543	S_Shore
cg10003262	chr17	1106589		chr17:1106812-1107379	N_Shore
cg00466249	chr17	79269884	*SLC38A10*	chr17:79268367-79269467	S_Shore
cg27069188	chr16	285785	*ITFG3*	chr16:284478-285395	S_Shore
cg01912040	chr17	1106553		chr17:1106812-1107379	N_Shore
cg13809441	chr6	6737631			OpenSea
cg12817436	chr19	1068561	*HMHA1*	chr19:1068438-1068764	Island
cg00131958	chr5	134073552	*CAMLG*	chr5:134074083-134074494	N_Shore
cg12845975	chr3	62305226	*C3orf14*	chr3:62304514-62304780	S_Shore
cg11380624	chr12	56222881	*DNAJC14*	chr12:56223263-56223598	N_Shore
cg01450228	chr16	27095462			OpenSea
cg06376715	chr1	3634918	*TP73*	chr1:3634633-3635101	Island
cg04041942	chr8	28747613	*INTS9*	chr8:28748070-28748431	N_Shore
cg11718235	chr19	56915184	*ZNF583*	chr19:56915356-56915856	N_Shore
cg02849695	chr1	159869960	*CCDC19*	chr1:159869901-159870143	Island
cg04870460	chr1	146644887	*PRKAB2*	chr1:146643470-146644501	S_Shore
cg09851528	chr11	117778029	*TMPRSS13*		OpenSea
cg08479073	chr1	207038584	*IL20*		OpenSea
cg03547757	chr20	57425515	*GNASAS*	chr20:57426729-57427047	N_Shore
cg13316433	chr5	83140500			OpenSea
cg23634348	chr4	967822	*DGKQ*	chr4:966732-968137	Island
cg07830134	chr16	1249204	*CACNA1H*	chr16:1248853-1249446	Island
cg02604095	chr9	32955790		chr9:32955651-32956278	Island
cg17379860	chr11	61159602	*TMEM216*	chr11:61159836-61160285	N_Shore
cg20016914	chr14	69263267	*C14orf181*	chr14:69259612-69262900	S_Shore

cg23061114	chr2	95540265	*TEKT4*	chr2:95540181-95540715	Island
cg18591727	chr10	8277320			OpenSea
cg00075920	chr5	177026190	*B4GALT7*	chr5:177026715-177027519	N_Shore
cg24011500	chr11	61049230	*VWCE*		OpenSea
cg08044427	chr2	20442131			OpenSea
cg13909895	chr22	51066142	*ARSA*	chr22:51066107-51067051	Island
cg03506502	chr4	174443567		chr4:174443365-174443948	Island
cg17164520	chr19	49540241	*CGB1*		OpenSea
cg16316624	chr11	61159649	*TMEM216*	chr11:61159836-61160285	N_Shore
cg27554551	chr16	4421486	*VASN*	chr16:4421642-4422240	N_Shore
cg12974393	chr11	107779099			OpenSea
cg03647724	chr16	10480289		chr16:10479686-10480254	S_Shore
cg04078980	chr1	228291486	*C1orf35*	chr1:228289669-228291184	S_Shore
cg14820927	chr18	34408166	*C18orf10*	chr18:34408558-34409289	N_Shore
cg08209724	chr17	30677138	*MIR632*	chr17:30677001-30677460	Island
cg12873476	chr8	142402728		chr8:142401533-142402494	S_Shore
cg20758593	chr5	169659863	*C5orf58*	chr5:169659830-169660229	Island
cg13850825	chr18	75998893		chr18:75998674-75999123	Island
cg11893552	chr6	168045268			OpenSea
cg21196136	chr19	48432069	*SNAR-C1*		OpenSea
cg19327213	chr19	47988783	*KPTN*	chr19:47987329-47987622	S_Shore
cg00142402	chr17	75956190		chr17:75953890-75955952	S_Shore
cg08378220	chr11	102678779			OpenSea
cg10592521	chr19	1560658	*MEX3D*	chr19:1555539-1556866	S_Shelf
cg13221347	chr10	134062614	*STK32C*		OpenSea
cg12078229	chr19	19245687	*TMEM161A*	chr19:19249182-19249443	N_Shelf
cg21218093	chr20	44600160	*ZNF335*	chr20:44599805-44601207	Island
cg26581886	chr22	17090862	*psiTPTE22*	chr22:17090455-17091224	Island
cg09313122	chr11	27827916			OpenSea
cg12493439	chr17	22203419			OpenSea
cg00093095	chr17	76481380	*DNAH17*		OpenSea
cg11510523	chr19	12306136		chr19:12305552-12306304	Island
cg04517524	chr14	94405342	*ASB2*	chr14:94405235-94406170	Island
cg22256960	chr15	77711686		chr15:77711625-77712910	Island
cg23934295	chr14	85995655	*FLRT2*	chr14:85996494-85996958	N_Shore
cg01435791	chr19	4585323		chr19:4584182-4585172	S_Shore
cg00478479	chr19	40904851	*PRX*	chr19:40904444-40904955	Island
cg01267709	chr14	23623756	*SLC7A8*		OpenSea
cg20306694	chr12	51718251	*BIN2*		OpenSea
cg05043349	chr8	65290113		chr8:65290108-65290946	Island
cg24346124	chr19	49540073	*CGB1*		OpenSea
cg17241310	chr1	91182856	*BARHL2*	chr1:91182509-91182857	Island
cg22790377	chr1	228291668	*C1orf35*	chr1:228289669-228291184	S_Shore
cg25294185	chr11	65487814	*RNASEH2C*	chr11:65487660-65488348	Island
cg03731646	chr18	74499372		chr18:74499316-74499641	Island
cg06960461	chr12	132950559			OpenSea
cg06026769	chr12	20704492	*PDE3A*	chr12:20704525-20706004	N_Shore
cg27257453	chr16	84096458	*MBTPS1*		OpenSea
cg11216554	chr10	124638983	*LOC399815*	chr10:124638743-124639793	Island
cg03503812	chr6	27357539	*ZNF391*	chr6:27356492-27357020	S_Shore
cg07229500	chr1	143467252		chr1:143467293-143467675	N_Shore
cg00689340	chr2	74668672	*RTKN*	chr2:74667593-74669403	Island
cg03662997	chr11	15963027		chr11:15962882-15963223	Island
cg00664205	chr17	77397758	*HRNBP3*		OpenSea
cg17762073	chr6	34024220	*GRM4*	chr6:34024201-34024457	Island
cg13287964	chr18	77905751	*LOC100130522*	chr18:77905297-77905566	S_Shore
cg13050716	chr3	127416580	*MGLL*		OpenSea
cg20713174	chr2	239048352	*KLHL30*	chr2:239048071-239048454	Island
cg17520215	chr12	100967016	*GAS2L3*	chr12:100967293-100967845	N_Shore

cg25397663	chr4	124622	*ZNF718*	chr4:124332-124841	Island
cg14392652	chr17	6927424	*BCL6B*	chr17:6925400-6927527	Island
cg16481667	chr11	70023578	*ANO1*	chr11:70023057-70023304	S_Shore
cg26649251	chr19	44598564	*ZNF224*	chr19:44598374-44598854	Island
cg12125332	chr14	75536561	*FAM164C*	chr14:75535825-75536376	S_Shore
cg25963505	chr7	135413023	*SLC13A4*		OpenSea
cg04584103	chr14	94405267	*ASB2*	chr14:94405235-94406170	Island
cg06663305	chr17	8095813		chr17:8092418-8093706	S_Shelf
cg20353001	chr6	1524618		chr6:1524153-1524543	S_Shore
cg20777656	chr15	22095674		chr15:22094182-22095519	S_Shore
cg07903023	chr2	239048072	*KLHL30*	chr2:239048071-239048454	Island
cg18363918	chr19	51829984	*IGLON5*	chr19:51830064-51831145	N_Shore
cg15109207	chr19	48614773	*PLA2G4C*		OpenSea
cg12497786	chr6	3849577	*FAM50B*	chr6:3849271-3851048	Island
cg01993818	chr14	34270437	*NPAS3*	chr14:34268946-34270438	Island
cg21923770	chr17	1183159	*TUSC5*		OpenSea
cg14590369	chr19	18113585	*ARRDC2*	chr19:18111566-18112481	S_Shore
cg11734329	chr19	7460671	*ARHGEF18*	chr19:7459779-7460376	S_Shore
cg12135344	chr1	6188120	*CHD5*	chr1:6187743-6188235	Island
cg02601352	chr8	21640100	*GFRA2*	chr8:21640161-21640365	N_Shore
cg06289138	chr6	33561186	*C6orf227*	chr6:33560892-33561189	Island
cg01621764	chr19	8809175	*ACTL9*	chr19:8807813-8808794	S_Shore
cg00201779	chr6	72131020	*C6orf155*	chr6:72129508-72130756	S_Shore
cg21695020	chr16	68014626	*DPEP3*	chr16:68013967-68014545	S_Shore
cg08124399	chr6	74104868	*DDX43*	chr6:74104425-74104878	Island
cg10631373	chr1	89457642	*RBMXL1*	chr1:89458034-89458604	N_Shore
cg01534423	chr17	18965556		chr17:18965477-18965728	Island
cg17395619	chr19	4607540		chr19:4607079-4607350	S_Shore
cg20678665	chr21	15096010		chr21:15095651-15096479	Island
cg16054907	chr22	50706065	*MAPK11*	chr22:50705406-50706553	Island
cg01517680	chr16	49499006		chr16:49498889-49499405	Island
cg19839825	chr19	7616011	*PNPLA6*	chr19:7614785-7615998	S_Shore
cg22478447	chr2	132588170		chr2:132585112-132588283	Island
cg14264194	chr1	211652549	*RD3*	chr1:211652306-211652754	Island
cg27009812	chr16	3062597	*CLDN9*	chr16:3063679-3063981	N_Shore
cg01969701	chr11	69286145		chr11:69285724-69286479	Island
cg04752871	chr2	121412432		chr2:121411983-121412315	S_Shore
cg05816193	chr6	26018127	*HIST1H1A*	chr6:26020671-26021125	N_Shelf
cg10197057	chr1	42612518		chr1:42611485-42611691	S_Shore
cg14836839	chr17	42706084			OpenSea
cg01781662	chr9	132387805	*METTL11A*	chr9:132388344-132388748	N_Shore
cg19786602	chr17	7966326			OpenSea
cg25433316	chr2	10197543	*CYS1*		OpenSea
cg22807681	chr5	174622933			OpenSea
cg14711997	chr7	150652864	*KCNH2*	chr7:150652807-150653080	Island
cg25836232	chr18	12306837	*TUBB6*	chr18:12307554-12309043	N_Shore
cg06501295	chr6	147523200		chr6:147525374-147525848	N_Shelf
cg19846667	chr14	22191856			OpenSea
cg27357306	chr10	63657059			OpenSea
cg01601573	chr4	144256835	*GAB1*	chr4:144256931-144258527	N_Shore
cg26537280	chr1	95699037	*RWDD3*	chr1:95699725-95700142	N_Shore
cg17277199	chr2	24397845	*C2orf84*	chr2:24397645-24398194	Island
cg01812146	chr1	94704178	*ARHGAP29*	chr1:94702690-94703344	S_Shore
cg04800569	chr19	55888003	*TMEM190*	chr19:55889005-55889476	N_Shore
cg20638675	chr10	45374906			OpenSea
cg27384088	chr6	72298275		chr6:72298274-72298528	Island
cg12560020	chr15	93823645			OpenSea
cg06286618	chr16	86987702			OpenSea
cg13451000	chr11	60383256	*C11orf64*		OpenSea

cg02566259	chr4	3043199		chr4:3043018-3043547	Island
cg20949342	chr13	20969084		chr13:20968851-20969220	Island
cg21198712	chr9	98534327			OpenSea
cg23008871	chr17	45949743		chr17:45949676-45949885	Island
cg12243133	chr11	66823616	*RHOD*	chr11:66824118-66824621	N_Shore
cg10697737	chr19	1673498		chr19:1671995-1672574	S_Shore
cg09362608	chr4	6659805		chr4:6660372-6660817	N_Shore
cg00337187	chr6	3054716			OpenSea
cg16844333	chr10	89167457			OpenSea
cg00469015	chr15	69744684	*RPLP1*	chr15:69745049-69745746	N_Shore
cg00516481	chr21	44073202	*PDE9A*	chr21:44073201-44074650	Island
cg24254488	chr7	25991654		chr7:25990012-25991320	S_Shore
cg02168442	chr8	102236522		chr8:102236153-102236949	Island
cg17181653	chr6	144417537	*SF3B5*	chr6:144416219-144416854	S_Shore
cg24168538	chr4	35527016			OpenSea
cg25594486	chr19	51165441	*SHANK1*	chr19:51165249-51165775	Island
cg23006519	chr2	121223740	*LOC84931*		OpenSea
cg01120509	chr16	698230	*WDR90*	chr16:698826-700244	N_Shore
cg09465703	chr16	733842	*JMJD8*	chr16:729438-735815	Island
cg22005401	chr19	18543829	*SSBP4*	chr19:18543828-18549161	Island
cg09663193	chr17	7486821	*MPDU1*	chr17:7486285-7487431	Island
cg06931941	chr17	7486770	*MPDU1*	chr17:7486285-7487431	Island
cg11874331	chr17	7486562	*MPDU1*	chr17:7486285-7487431	Island
cg12244647	chr22	50737978	*PLXNB2*	chr22:50737809-50738961	Island
cg13546736	chr17	7486858	*MPDU1*	chr17:7486285-7487431	Island
cg02617560	chr8	32785757			OpenSea
cg03521774	chr8	81478172			OpenSea
cg06653052	chr19	1275167	*C19orf24*	chr19:1274469-1276358	Island
cg02482603	chr1	174843754	*RABGAP1L*		OpenSea
cg16009311	chr8	119086762	*EXT1*		OpenSea
cg06833564	chr16	4665307	*FAM100A*	chr16:4664006-4666782	Island
cg18033770	chr9	139259074	*CARD9*	chr9:139257842-139259075	Island
cg03871267	chr5	67483475			OpenSea
cg16856833	chr3	87138640		chr3:87138225-87138546	S_Shore
cg11980800	chr1	9748107	*PIK3CD*	chr1:9749295-9750228	N_Shore
cg21330831	chr6	33283092	*ZBTB22*	chr6:33282855-33283196	Island
cg02213139	chr1	85527754	*WDR63*		OpenSea
cg03870261	chr19	2428514	*LMNB2*	chr19:2428349-2428731	Island
cg10955566	chr1	1003126		chr1:1002663-1005318	Island
cg09459982	chr6	33386556	*CUTA*	chr6:33385678-33386262	S_Shore
cg21653641	chr19	49222892	*MAMSTR*	chr19:49223054-49223312	N_Shore
cg13540341	chr19	49222985	*MAMSTR*	chr19:49223054-49223312	N_Shore
cg03464655	chr2	165811982	*SLC38A11*		OpenSea
cg06505619	chr16	698072	*WDR90*	chr16:698826-700244	N_Shore
cg10404009	chr19	49223845	*MAMSTR*	chr19:49223833-49224276	Island
cg10599948	chr19	49244571	*IZUMO1*	chr19:49242018-49242962	S_Shore
cg01607625	chr6	32847830	*PPP1R2P1*	chr6:32847498-32847846	Island
cg26343883	chr2	164205032		chr2:164204740-164205039	Island
cg22987448	chr19	8591364	*MYO1F*	chr19:8591294-8591842	Island
cg23995446	chr19	7934807		chr19:7933263-7934898	Island
cg18804667	chr5	58883392	*PDE4D*		OpenSea
cg09413013	chr17	76137024	*TMC8*	chr17:76136513-76137184	Island
cg27475923	chr11	63530997	*C11orf95*	chr11:63529888-63531834	Island
cg13526264	chr6	113886304			OpenSea
cg16404106	chr11	86085623	*CCDC81*	chr11:86085693-86086006	N_Shore
cg19878214	chr5	70742551			OpenSea
cg27454064	chr12	64215611		chr12:64215499-64215990	Island
cg11334165	chr5	145758782		chr5:145758569-145758882	Island
cg01856645	chr5	78365647	*DMGDH*	chr5:78365298-78365711	Island

cg05358404	chr20	62288950	*RTEL1*	chr20:62288941-62290058	Island
cg27576241	chr1	155034570	*ADAM15*	chr1:155034925-155036578	N_Shore
cg22198044	chr4	2819614	*SH3BP2*	chr4:2819499-2820429	Island
cg15967501	chr19	49143404	*SEC1*	chr19:49140156-49141131	S_Shelf
cg26209676	chr19	56154382	*ZNF581*	chr19:56154041-56154405	Island
cg04181892	chr2	54936027			OpenSea
cg16717549	chr21	45705699	*AIRE*	chr21:45705428-45706044	Island
cg13069100	chr1	155035008	*ADAM15*	chr1:155034925-155036578	Island
cg20734996	chr17	7234027	*NEURL4*	chr17:7232330-7233151	S_Shore
cg21399832	chr4	117847410		chr4:117847087-117847522	Island
cg07716847	chr16	67233549	*ELMO3*	chr16:67233032-67233862	Island
cg12176783	chr20	62694000	*TCEA2*	chr20:62693942-62695282	Island
cg12082025	chr19	1064218	*ABCA7*	chr19:1063544-1064265	Island
cg08941173	chr11	125034971	*PKNOX2*	chr11:125034896-125036426	Island
cg27168291	chr6	33282997	*ZBTB22*	chr6:33282855-33283196	Island
cg11921736	chr3	113160437	*WDR52*	chr3:113160299-113160641	Island
cg21235532	chr18	19476870		chr18:19476852-19477103	Island
cg02472291	chr8	67783002	*C8orf45*		OpenSea
cg23891273	chr12	102036400	*MYBPC1*	chr12:102036253-102036461	Island
cg23378165	chr17	9550256	*USP43*	chr17:9548389-9549616	S_Shore
cg16356856	chr1	74663583	*LRRIQ3*	chr1:74663826-74664035	N_Shore
cg16415457	chr3	23713441			OpenSea
cg24602417	chr8	102236283		chr8:102236153-102236949	Island
cg08122232	chr17	74100132	*EXOC7*	chr17:74099738-74100055	S_Shore
cg25043559	chr18	12948780	*SEH1L*	chr18:12947924-12948654	S_Shore
cg22040809	chr6	26522578	*HCG11*	chr6:26521947-26522579	Island
cg18605031	chr3	49314155	*C3orf62*	chr3:49314437-49314815	N_Shore
cg15352671	chr11	1331497	*LOC255512*	chr11:1330390-1331498	Island
cg20318748	chr20	25605178	*NANP*	chr20:25603815-25605187	Island
cg13951491	chr1	45793032	*HPDL*	chr1:45792419-45793301	Island
cg03614916	chr18	22007271	*IMPACT*	chr18:22006310-22007007	S_Shore
cg23725986	chr2	75938438	*C2orf3*	chr2:75937657-75938139	S_Shore
cg12924095	chr5	151150029	*G3BP1*	chr5:151150014-151152086	Island
cg06638811	chr9	137030555	*RNU6ATAC*	chr9:137028717-137030556	Island
cg08962798	chr5	179106333	*CBY3*	chr5:179105169-179106412	Island
cg06480265	chr6	24722060		chr6:24720387-24721920	S_Shore
cg09159285	chr1	6664134	*KLHL21*	chr1:6661776-6663844	S_Shore
cg14611683	chr1	45452580	*EIF2B3*	chr1:45452037-45452444	S_Shore
cg18407136	chr15	83681092	*C15orf40*	chr15:83680007-83680532	S_Shore
cg08602346	chr13	113242997	*TUBGCP3*	chr13:113241711-113241925	S_Shore
cg14593033	chr17	56296980	*MKS1*	chr17:56296488-56296842	S_Shore
cg12065943	chr17	19881925	*AKAP10*	chr17:19883325-19883610	N_Shore
cg23373640	chr14	65696480			OpenSea
cg05465916	chr17	7819762	*LOC284023*	chr17:7818938-7819763	Island
cg01039401	chr12	3070406	*TEAD4*	chr12:3067960-3069444	S_Shore
cg06975979	chr7	150020025	*LRRC61*	chr7:150019950-150020752	Island
cg21498471	chr19	33183713	*NUDT19*	chr19:33182609-33183562	S_Shore
cg18427905	chr2	232825958	*DIS3L2*	chr2:232826077-232826836	N_Shore
cg10045137	chr2	25383940	*POMC*	chr2:25383939-25384763	Island
cg10441379	chr5	72793693	*BTF3*	chr5:72794040-72794641	N_Shore
cg23915527	chr1	161368787		chr1:161368863-161369932	N_Shore
cg14870223	chr3	39093076	*WDR48*	chr3:39093469-39094025	N_Shore
cg06868955	chr12	54069197	*ATP5G2*	chr12:54069625-54070177	N_Shore
cg13430552	chr2	25427652		chr2:25427101-25427577	S_Shore
cg14576824	chr1	213224402	*RPS6KC1*	chr1:213224611-213224867	N_Shore
cg01404873	chr13	50701050	*DLEU2*	chr13:50697984-50702286	Island
cg06939851	chr1	226112534	*PYCR2*	chr1:226111642-226112552	Island
cg04220541	chr8	48172992	*KIAA0146*	chr8:48172791-48172993	Island
cg10105237	chr8	135726091	*ZFAT*	chr8:135724770-135725552	S_Shore

cg26751513	chr16	1542910	*TELO2*	chr16:1543098-1544426	N_Shore
cg24435209	chr22	17640972	*CECR5*	chr22:17639481-17640714	S_Shore
cg18829411	chr21	40722023	*HMGN1*	chr21:40720185-40721625	S_Shore
cg12265130	chr11	59437344	*PATL1*	chr11:59436201-59437232	S_Shore
cg12151328	chr3	112709521	*GTPBP8*	chr3:112709666-112710270	N_Shore
cg14959908	chr20	2736884	*EBF4*	chr20:2732746-2733630	S_Shelf
cg03456393	chr3	193310565	*OPA1*	chr3:193310824-193311188	N_Shore
cg03804621	chr10	124638756	*FAM24B*	chr10:124638743-124639793	Island
cg26312542	chr11	112038104	*TEX12*		OpenSea
cg00373499	chr5	137224723	*PKD2L2*	chr5:137224986-137225477	N_Shore
cg04861271	chr14	23789531	*PABPN1*	chr14:23790212-23790659	N_Shore
cg05072008	chr7	50518647	*FIGNL1*	chr7:50517627-50518668	Island
cg08558340	chr7	100472263	*SRRT*	chr7:100472375-100473393	N_Shore
cg15661865	chr2	27632926	*PPM1G*	chr2:27631884-27632822	S_Shore
cg13028113	chr5	60458778	*C5orf43*	chr5:60457840-60458654	S_Shore
cg14307853	chr9	91924749	*CKS2*	chr9:91925372-91926639	N_Shore
cg10432947	chr4	2062441	*NAT8L*	chr4:2059928-2063181	Island
cg17055585	chr5	25190850		chr5:25190503-25191113	Island
cg22993195	chr3	49756301	*AMIGO3*	chr3:49755763-49757317	Island
cg05868531	chr2	232348602		chr2:232348216-232348866	Island
cg22189786	chr22	42395067	*WBP2NL*	chr22:42394589-42395255	Island
cg16740427	chr22	42469976	*FAM109B*	chr22:42470035-42470669	N_Shore
cg25825627	chr12	127631198		chr12:127630511-127631199	Island
cg14625636	chr9	100000026	*KIAA1529*	chr9:100000463-100000820	N_Shore
cg01290153	chr19	55987415	*ZNF628*	chr19:55987544-55988171	N_Shore
cg26440467	chr8	145654605	*VPS28*	chr8:145654517-145654810	Island
cg16245698	chr18	43355989		chr18:43355318-43355698	S_Shore
cg04036182	chr15	45458818		chr15:45458801-45459580	Island
cg11367159	chr12	110044531			OpenSea
cg05950212	chr1	149860711	*HIST2H2AB*	chr1:149857769-149859470	S_Shore
cg05742863	chr22	45809319	*SMC1B*	chr22:45809191-45809953	Island
cg17545182	chr1	247694041	*LOC148824*	chr1:247694035-247694501	Island
cg12403142	chr1	92012408		chr1:92012336-92012656	Island
cg20744362	chr22	50050164	*C22orf34*		OpenSea
cg11724970	chr7	27182493	*HOXA5*	chr7:27182613-27185562	N_Shore
cg20974609	chr7	27181671	*HOXA5*	chr7:27182613-27185562	N_Shore
cg23454797	chr7	27183990	*HOXA5*	chr7:27182613-27185562	Island
cg09343092	chr7	27185282	*HOXA6*	chr7:27182613-27185562	Island
cg03368099	chr7	27184521	*HOXA5*	chr7:27182613-27185562	Island
cg17432857	chr7	27184438	*HOXA5*	chr7:27182613-27185562	Island
cg20517050	chr7	27183806	*HOXA5*	chr7:27182613-27185562	Island
cg05928186	chr7	27187102	*HOXA6*	chr7:27186927-27187692	Island
cg23129930	chr7	27186993	*HOXA6*	chr7:27186927-27187692	Island
cg05579037	chr7	27184853		chr7:27182613-27185562	Island
cg09650488	chr7	1687007		chr7:1686716-1687008	Island
cg27151303	chr7	27184821		chr7:27182613-27185562	Island
cg14853974	chr17	8218887	*ARHGEF15*		OpenSea
cg04635635	chr15	30163726			OpenSea
cg02341556	chr11	118781978	*BCL9L*	chr11:118781060-118781732	S_Shore
cg01687997	chr7	1083072	*C7orf50*	chr7:1084387-1084595	N_Shore
cg04265576	chr7	27187372	*HOXA6*	chr7:27186927-27187692	Island
cg07278425	chr7	27137922		chr7:27135342-27136736	S_Shore
cg23256343	chr15	30163825			OpenSea
cg05612279	chr12	52241471		chr12:52240772-52242547	Island
cg10601100	chr3	64085148	*PRICKLE2*		OpenSea
cg03529432	chr7	27187502	*HOXA6*	chr7:27186927-27187692	Island
cg24460369	chr12	109549041		chr12:109548899-109549174	Island
cg07441944	chr11	44623459	*CD82*		OpenSea
cg24168308	chr21	27372387	*APP*		OpenSea

cg08319974	chr6	164506861			OpenSea
cg20370909	chr7	4859229	*RADIL*	chr7:4855990-4856212	S_Shelf
cg17461234	chr12	133373022	*GOLGA3*	chr12:133369195-133370290	S_Shelf
cg04226804	chr1	16556249		chr1:16553326-16553660	S_Shelf
cg25694032	chr6	164507164			OpenSea
cg11246938	chr12	115112433	*TBX3*	chr12:115112014-115112627	Island
cg08380478	chr10	132683454			OpenSea
cg12444411	chr7	2802554	*GNA12*	chr7:2802440-2802698	Island
cg20710519	chr5	6775909			OpenSea
cg03607916	chr2	58478696		chr2:58478734-58478938	N_Shore
cg10651121	chr12	109549097		chr12:109548899-109549174	Island
cg04053108	chr12	6166028	*VWF*	chr12:6165941-6166257	Island
cg10482201	chr16	3493423	*ZNF597*	chr16:3493098-3493569	Island
cg24095222	chr5	161178787			OpenSea
cg16278496	chr14	98444476	*C14orf64*		OpenSea
cg00232832	chr15	30163660			OpenSea
cg27365295	chr16	1483170		chr16:1482551-1483201	Island
cg08959039	chr13	111062266	*COL4A2*		OpenSea
cg13713522	chr8	141109051	*TRAPPC9*	chr8:141107837-141110984	Island
cg11874426	chr11	118781731	*BCL9L*	chr11:118781060-118781732	Island
cg23496597	chr20	57463725	*GNAS*	chr20:57463652-57467739	Island
cg22469274	chr7	27187553	*HOXA6*	chr7:27186927-27187692	Island
cg23355997	chr17	7154831	*DULLARD*	chr17:7154424-7156126	Island
cg19163395	chr17	42164433	*HDAC5*		OpenSea
cg07593523	chr5	58082800	*RAB3C*		OpenSea
cg19254235	chr13	48895173	*RB1*	chr13:48892635-48893857	S_Shore
cg09954631	chr9	94648732	*ROR2*	chr9:94648653-94648964	Island
cg22872349	chr15	81410745		chr15:81410715-81411067	Island
cg00383793	chr3	50488162	*CACNA2D2*		OpenSea
cg14660024	chr17	7154518	*C17orf81*	chr17:7154424-7156126	Island
cg06388363	chr6	164507305			OpenSea
cg02088785	chr13	99223336	*STK24*		OpenSea
cg03906434	chr7	27231819		chr7:27231805-27233097	Island
cg03395546	chr19	41222599	*ADCK4*	chr19:41222598-41223162	Island
cg22755142	chr2	16243377			OpenSea
cg06495710	chr2	167449521			OpenSea
cg14160422	chr19	10735548	*SLC44A2*	chr19:10735999-10736396	N_Shore
cg23712970	chr14	23540735	*ACIN1*		OpenSea
cg15942979	chr8	22497061	*BIN3*		OpenSea
cg02324835	chr8	28919075			OpenSea
cg14614581	chr10	73844130	*SPOCK2*	chr10:73846806-73848233	N_Shelf
cg01156816	chr4	132984152			OpenSea
cg22040889	chr11	64073594	*ESRRA*	chr11:64072064-64073914	Island
cg01748573	chr20	57463530	*GNAS*	chr20:57463652-57467739	N_Shore
cg01320698	chr1	9788517	*PIK3CD*	chr1:9790292-9790704	N_Shore
cg16522260	chr11	114958426			OpenSea
cg20337103	chr11	118781778	*BCL9L*	chr11:118781060-118781732	S_Shore
cg01368900	chr3	13113959	*IQSEC1*	chr3:13114627-13115245	N_Shore
cg23198262	chr14	22978073			OpenSea
cg13138954	chr15	22546672		chr15:22545218-22547291	Island
cg22989533	chr13	82264223			OpenSea
cg10543953	chr15	35529612	*LOC723972*		OpenSea
cg02478023	chr19	57351322	*MIMT1*	chr19:57351283-57351995	Island
cg16059991	chr15	99975470			OpenSea
cg07659054	chr7	27134225	*HOXA1*	chr7:27134097-27134303	Island
cg09677938	chr13	23423308		chr13:23422687-23422896	S_Shore
cg08867825	chr1	102312608	*OLFM3*		OpenSea
cg05637296	chr17	76129475	*TMC8*	chr17:76127521-76128406	S_Shore
cg01525879	chr11	67207498	*CORO1B*	chr11:67203319-67203662	S_Shelf

cg27376514	chr17	17058422	*MPRIP*	chr17:17061857-17062186	N_Shelf
cg18751141	chr7	27138173		chr7:27135342-27136736	S_Shore
cg10046893	chr14	101456920	*SNORD114-30*		OpenSea
cg08884752	chr1	2162001	*SKI*	chr1:2158212-2161173	S_Shore
cg03431524	chr7	100142441	*AGFG2*		OpenSea
cg20918563	chr3	196263520			OpenSea
cg07512258	chr10	75619420	*CAMK2G*		OpenSea
cg14192130	chr6	167535764	*CCR6*		OpenSea
cg21417130	chr1	19400922			OpenSea
cg14506192	chr12	122711988	*DIABLO*	chr12:122710054-122711371	S_Shore
cg20988098	chr6	157931791	*ZDHHC14*		OpenSea
cg07757252	chr3	52274561	*TWF2*	chr3:52272839-52273631	S_Shore
cg00537673	chr6	43141823	*SRF*	chr6:43142013-43142217	N_Shore
cg02340818	chr8	145808436	*KIAA1688*	chr8:145806258-145806713	S_Shore
cg16265553	chr1	226911669	*ITPKB*		OpenSea
cg22471075	chr22	49765229		chr22:49764995-49765272	Island
cg26374305	chr14	101293856	*MEG3*	chr14:101293427-101294433	Island
cg16880946	chr7	27188465	*HOXA6*	chr7:27190274-27191115	N_Shore
cg22859054	chr6	32847548	*PPP1R2P1*	chr6:32847498-32847846	Island
cg27500647	chr19	51602230	*CTU1*	chr19:51601822-51602260	Island
cg15655154	chr3	113604241	*GRAMD1C*	chr3:113603863-113604246	Island
cg08446215	chr11	2721366	*KCNQ1OT1*	chr11:2720410-2722087	Island
cg01010328	chr1	156551873	*TTC24*		OpenSea
cg01424562	chr14	69256677	*ZFP36L1*	chr14:69256676-69257036	Island
cg14179389	chr1	92947961	*GFI1*	chr1:92945907-92952609	Island
cg21106946	chr13	65533500			OpenSea
cg05346902	chr19	47910374	*MEIS3*	chr19:47910107-47910563	Island
cg08145373	chr11	2407008	*CD81*	chr11:2406711-2406927	S_Shore
cg06313676	chr16	87978974			OpenSea
cg24437859	chr12	7066614	*PTPN6*		OpenSea
cg05321594	chr8	330175			OpenSea
cg02317251	chr4	10116515	*WDR1*	chr4:10117575-10118952	N_Shore
cg21705506	chr17	33842181	*SLFN12L*		OpenSea
cg25550279	chr7	53254983		chr7:53254982-53255204	Island
cg17220933	chr1	206729613	*RASSF5*	chr1:206730397-206730908	N_Shore
cg21463790	chr7	14025907	*ETV1*	chr7:14029378-14029593	N_Shelf
cg04042333	chr17	1104665		chr17:1106812-1107379	N_Shelf
cg20062691	chr1	949392	*ISG15*	chr1:949329-949851	Island
cg18138400	chr11	64073849	*ESRRA*	chr11:64072064-64073914	Island
cg07899411	chr12	123604434			OpenSea
cg23511824	chr14	54816172			OpenSea
cg06666376	chr19	3480596	*C19orf77*	chr19:3481782-3482048	N_Shore
cg19655952	chr7	114055204	*FOXP2*		OpenSea
cg08659421	chr16	3117862	*IL32*		OpenSea
cg14658607	chr2	139660341		chr2:139660186-139660395	Island
cg09016797	chr7	155790956			OpenSea
cg08169020	chr14	69256888	*ZFP36L1*	chr14:69256676-69257036	Island
cg09107344	chr12	53612734	*RARG*	chr12:53613716-53615103	N_Shore
cg25198255	chr10	134878094		chr10:134876330-134878539	Island
cg17995557	chr10	126289971	*LHPP*		OpenSea
cg17282688	chr1	11516947			OpenSea
cg00367649	chr3	66702975			OpenSea
cg21565496	chr13	40762150			OpenSea
cg17001153	chr1	206742216	*RASSF5*		OpenSea
cg08883853	chr7	1026168	*CYP2W1*	chr7:1026167-1026395	Island
cg02519218	chr12	133424286	*CHFR*		OpenSea
cg23901967	chr17	36890321	*CISD3*	chr17:36886108-36886525	S_Shelf
cg00106345	chr7	27138396		chr7:27135342-27136736	S_Shore
cg23629187	chr8	125576109	*MTSS1*		OpenSea

cg08196561	chr16	87525539	*ZCCHC14*		OpenSea
cg00689612	chr17	76899627	*TIMP2*	chr17:76897078-76897298	S_Shelf
cg22459924	chr19	2607850	*GNG7*	chr19:2607725-2607967	Island
cg12756527	chr10	11284548	*CUGBP2*		OpenSea
cg05066096	chr7	4781872	*FOXK1*	chr7:4784820-4785058	N_Shelf
cg01490040	chr2	151399444			OpenSea
cg10782923	chr12	56329731	*DGKA*	chr12:56329518-56329790	Island
cg11374834	chr3	75263691			OpenSea
cg15526535	chr1	12238546	*TNFRSF1B*		OpenSea
cg17922695	chr17	75451809	*42987*	chr17:75447477-75447821	S_Shelf
cg01246390	chr12	9392651		chr12:9392459-9392686	Island
cg16222326	chr8	1708438		chr8:1711234-1712654	N_Shelf
cg11733958	chr17	17023057	*MPRIP*		OpenSea
cg02082843	chr14	92339595	*FBLN5*		OpenSea
cg16265859	chr1	19401069	*UBR4*		OpenSea
cg26103719	chr14	92336250	*FBLN5*		OpenSea
cg02668248	chr19	16437789	*KLF2*	chr19:16435202-16438064	Island
cg06201717	chr1	2949633			OpenSea
cg05932042	chr17	40822653	*PLEKHH3*	chr17:40821935-40823573	Island
cg16898066	chr6	25726437	*HIST1H2BA*	chr6:25726454-25726722	N_Shore
cg05541267	chr3	50647374	*CISH*	chr3:50648939-50649829	N_Shore
cg13390570	chr5	1255616	*TERT*	chr5:1256968-1257222	N_Shore
cg21422623	chr1	226912213	*ITPKB*		OpenSea
cg02640489	chr2	219153935	*PNKD*	chr2:219156977-219157380	N_Shelf
cg25526001	chr10	85939451	*C10orf99*		OpenSea
cg22241838	chr11	60680006	*TMEM109*	chr11:60681552-60682268	N_Shore
cg16349277	chr4	83323642			OpenSea
cg17972213	chr1	101704898	*S1PR1*	chr1:101701785-101702068	S_Shelf
cg21918313	chr2	73496203	*FBXO41*	chr2:73495920-73496910	Island
cg24231380	chr2	242813914	*C2orf85*	chr2:242815056-242815599	N_Shore
cg08978665	chr16	3115707	*IL32*		OpenSea
cg16235962	chr11	118754507	*CXCR5*		OpenSea
cg00589791	chr5	30429329			OpenSea
cg25382214	chr1	3105252	*PRDM16*	chr1:3104956-3105253	Island
cg14843920	chr17	75451932	*42987*		OpenSea
cg22572362	chr5	501938	*SLC9A3*	chr5:501835-502075	Island
cg01984743	chr5	172751331	*STC2*	chr5:172754056-172757098	N_Shelf
cg15607708	chr19	54041308	*ZNF331*	chr19:54040812-54041857	Island
cg15079934	chr11	1892307	*LSP1*	chr11:1892037-1892556	Island
cg03690763	chr11	133734501		chr11:133734701-133735164	N_Shore
cg25278298	chr4	1305282	*MAEA*	chr4:1303490-1303835	S_Shore
cg26146569	chr15	31637592	*KLF13*		OpenSea
cg08059678	chr16	70728425	*VAC14*		OpenSea
cg04387073	chr6	165415258			OpenSea
cg04252928	chr16	89168599	*ACSF3*	chr16:89167116-89167482	S_Shore
cg07692540	chr3	38244838	*OXSR1*		OpenSea
cg04086531	chr16	49732376	*ZNF423*	chr16:49732348-49732712	Island
cg12172478	chr3	16357591	*RFTN1*		OpenSea
cg01757168	chr3	128215565		chr3:128215212-128216905	Island
cg00357958	chr12	123215010	*GPR81*		OpenSea
cg02231590	chr2	231737958	*ITM2C*		OpenSea
cg08039560	chr21	45575832		chr21:45575451-45575833	Island
cg15339435	chr11	128169380			OpenSea
cg11525252	chr1	158041312	*KIRREL*		OpenSea
cg09312904	chr1	117281442		chr1:117284383-117284627	N_Shelf
cg00517080	chr11	134098583	*VPS26B*	chr11:134093696-134095491	S_Shelf
cg09117448	chr16	31439393	*COX6A2*	chr16:31438891-31439682	Island
cg06517940	chr1	32717620	*LCK*	chr1:32713918-32714186	S_Shelf
cg26888227	chr10	45697308		chr10:45693718-45694890	S_Shelf

cg25323888	chr6	11933336			OpenSea
cg07651189	chr4	88753434	*MEPE*		OpenSea
cg04819180	chr17	80829157	*TBCD*	chr17:80829533-80829871	N_Shore
cg00669330	chr2	242810917	*C2orf85*	chr2:242808405-242808618	S_Shelf
cg10963061	chr1	158147729		chr1:158147433-158147854	Island
cg01579052	chr16	788035	*NARFL*	chr16:787070-787327	S_Shore
cg19528654	chr11	67251154	*AIP*	chr11:67250344-67250864	S_Shore
cg20997320	chr21	47845610	*PCNT*	chr21:47845595-47845864	Island
cg01134260	chr7	63924733		chr7:63924400-63924741	Island
cg08109681	chr6	166825084	*RPS6KA2*	chr6:166825827-166826439	N_Shore
cg23479730	chr14	99681757	*BCL11B*		OpenSea
cg23892025	chr2	63849775	*LOC388955*	chr2:63849884-63850164	N_Shore
cg19752602	chr5	153039227	*GRIA1*		OpenSea
cg18349306	chr4	146267385			OpenSea
cg05595469	chr21	47845863	*PCNT*	chr21:47845595-47845864	Island
cg17881200	chr7	27138850		chr7:27135342-27136736	S_Shelf
cg02284273	chr10	14051679	*FRMD4A*		OpenSea
cg10015572	chr11	124071575			OpenSea
cg18350391	chr16	3115552	*IL32*		OpenSea
cg05437132	chr21	10990903	*TPTE*	chr21:10989913-10991413	Island
cg26912105	chr16	84553585			OpenSea
cg12444081	chr17	16521159		chr17:16520938-16521317	Island
cg12047941	chr11	315908		chr11:315739-316539	Island
cg09163005	chr12	14926986	*H2AFJ*	chr12:14927291-14928023	N_Shore
cg14690181	chr5	115387984			OpenSea
cg16239536	chr19	1079617	*HMHA1*	chr19:1079682-1081937	N_Shore
cg08418872	chr12	6442954	*TNFRSF1A*		OpenSea
cg24549912	chr5	50692281		chr5:50694984-50695352	N_Shelf
cg12077963	chr17	4079306	*ANKFY1*		OpenSea
cg23026906	chr1	207819196	*CR1L*	chr1:207818389-207818952	S_Shore
cg07499142	chr1	9788715	*PIK3CD*	chr1:9790292-9790704	N_Shore
cg15997393	chr7	1961869	*MAD1L1*		OpenSea
cg01123946	chr16	89163343	*ACSF3*	chr16:89167116-89167482	N_Shelf
cg14378925	chr11	64635728	*EHD1*		OpenSea
cg01945624	chr4	8230847	*SH3TC1*		OpenSea
cg14085952	chr9	139496844			OpenSea
cg12413156	chr20	62368256	*LIME1*	chr20:62368955-62371962	N_Shore
cg06329392	chr13	40762435			OpenSea
cg22022580	chr19	37266605		chr19:37266594-37267027	Island
cg18128536	chr17	47092178	*IGF2BP1*	chr17:47091973-47092211	Island
cg25607383	chr6	30853569	*DDR1*	chr6:30852102-30852676	S_Shore
cg11227278	chr2	23749277	*KLHL29*	chr2:23749086-23749291	Island
cg04582263	chr16	67973788	*LCAT*	chr16:67969290-67970418	S_Shelf
cg07970799	chr6	6614719	*LOC285780*		OpenSea
cg20513976	chr20	62367893	*LIME1*	chr20:62368955-62371962	N_Shore
cg13603599	chr11	67251939	*AIP*	chr11:67250344-67250864	S_Shore
cg16787483	chr13	88328251	*SLITRK5*	chr13:88329394-88329885	N_Shore
cg25225693	chr7	99816988	*GATS*		OpenSea
cg25598685	chr11	42617544			OpenSea
cg07637658	chr19	45887042	*PPP1R13L*	chr19:45888768-45889465	N_Shore
cg24741873	chr10	82223998	*TSPAN14*		OpenSea
cg18030943	chr3	182876556	*LAMP3*	chr3:182880277-182880722	N_Shelf
cg08006309	chr16	1587810	*IFT140*	chr16:1587683-1587899	Island
cg01154505	chr2	112940409	*FBLN7*	chr2:112939287-112939563	S_Shore
cg26472802	chr21	45713719	*AIRE*	chr21:45713509-45713813	Island
cg18338046	chr5	133452188	*TCF7*	chr5:133449320-133451260	S_Shore
cg13750061	chr11	118214845	*CD3G*		OpenSea
cg14163311	chr1	206729553	*RASSF5*	chr1:206730397-206730908	N_Shore
cg17709873	chr6	31540456	*LTA*		OpenSea

cg09985344	chr16	84648441	*COTL1*	chr16:84650960-84651565	N_Shelf
cg09232358	chr14	23015657			OpenSea
cg13298528	chr11	118763863	*CXCR5*		OpenSea
cg02119848	chr3	125826288	*ALDH1L1*		OpenSea
cg26348243	chr6	31540461	*LTA*		OpenSea
cg02324006	chr19	1080034	*HMHA1*	chr19:1079682-1081937	Island
cg26752663	chr2	25142016	*ADCY3*	chr2:25142472-25143689	N_Shore
cg20334591	chr4	39775714	*UBE2K*	chr4:39771679-39771944	S_Shelf
cg12033075	chr1	9788767	*PIK3CD*	chr1:9790292-9790704	N_Shore
cg21653586	chr11	10530636			OpenSea
cg24449463	chr1	168025552	*DCAF6*	chr1:168025508-168025722	Island
cg18799102	chr3	155760037			OpenSea
cg22514722	chr3	127473755	*MGLL*		OpenSea
cg06243989	chr15	25425615	*SNORD115-6*		OpenSea
cg21482472	chr13	81229358		chr13:81229047-81229445	Island
cg09549073	chr7	27183274	*HOXA5*	chr7:27182613-27185562	Island
cg14918744	chr8	23018464	*TNFRSF10D*	chr8:23021056-23021588	N_Shelf
cg24973755	chr4	1304972	*MAEA*	chr4:1304768-1305114	Island
cg02611934	chr13	88329407	*SLITRK5*	chr13:88329394-88329885	Island
cg20805133	chr2	242802192	*PDCD1*	chr2:242805752-242806034	N_Shelf
cg03363478	chr6	170268808			OpenSea
cg19529732	chr12	122712101	*DIABLO*	chr12:122710054-122711371	S_Shore
cg21673175	chr2	132202532	*LOC401010*	chr2:132202265-132203257	Island
cg23743554	chr11	65321226	*LTBP3*	chr11:65321225-65321823	Island
cg25657218	chr17	22017062		chr17:22016769-22017189	Island
cg16385865	chr12	132670412		chr12:132673993-132675241	N_Shelf
cg18825594	chr2	10184650	*KLF11*	chr2:10182635-10184702	Island
cg05163268	chr5	180116385			OpenSea
cg04520704	chr22	18325160	*MICAL3*	chr22:18324563-18324788	S_Shore
cg24449629	chr19	52646265		chr19:52642613-52642841	S_Shelf
cg12827637	chr14	69256791	*ZFP36L1*	chr14:69256676-69257036	Island
cg06357748	chr12	1025529	*RAD52*		OpenSea
cg04006327	chr11	72533487	*ATG16L2*	chr11:72532612-72533774	Island
cg09757087	chr16	1585644	*IFT140*	chr16:1583809-1584641	S_Shore
cg24616382	chr7	99767035	*GAL3ST4*	chr7:99768884-99769559	N_Shore
cg14356919	chr21	46890997	*COL18A1*	chr21:46890848-46891115	Island
cg16093065	chr22	40720633	*TNRC6B*		OpenSea
cg24331317	chr3	19017607			OpenSea
cg20793665	chr2	232549224		chr2:232545347-232546143	S_Shelf
cg24723883	chr19	2608495	*GNG7*	chr19:2607725-2607967	S_Shore
cg10325497	chr1	32739049	*LCK*	chr1:32740710-32741544	N_Shore
cg02963266	chr14	99681710	*BCL11B*		OpenSea
cg14957089	chr11	331092		chr11:330785-331110	Island
cg25569480	chr2	48844892	*GTF2A1L*	chr2:48844751-48844985	Island
cg19484299	chr2	234775274	*MSL3L2*	chr2:234776882-234777098	N_Shore
cg08329113	chr16	70771142	*VAC14*		OpenSea
cg08288818	chr15	55920079	*PRTG*		OpenSea
cg11430096	chr6	110968061	*CDK19*	chr6:110967011-110967823	S_Shore
cg17187762	chr22	28070120		chr22:28073920-28074244	N_Shelf
cg15739904	chr8	125313885		chr8:125313572-125314097	Island
cg08466385	chr6	18368294			OpenSea
cg01791634	chr17	76130139	*TMC8*	chr17:76127521-76128406	S_Shore
cg23939642	chr17	79225757	*SLC38A10*		OpenSea
cg11015251	chr7	27170554	*HOXA4*	chr7:27169572-27170638	Island
cg07114422	chr1	2183438	*SKI*		OpenSea
cg27583010	chr16	30198505	*CORO1A*	chr16:30194747-30195357	S_Shelf
cg26821758	chr2	55515029	*CCDC88A*		OpenSea
cg19905757	chr15	68924127	*CORO2B*		OpenSea
cg14457074	chr5	180408785			OpenSea

cg09390241	chr4	8174148			OpenSea
cg16272981	chr5	1489889	*LPCAT1*	chr5:1489799-1490071	Island
cg22726349	chr1	2990678	*PRDM16*	chr1:2990030-2990718	Island
cg14977069	chr20	62367698	*LIME1*	chr20:62368955-62371962	N_Shore
cg25417842	chr7	150822987	*AGAP3*	chr7:150822654-150823387	Island
cg19933320	chr7	64125401	*ZNF107*	chr7:64126715-64126994	N_Shore
cg08121102	chr1	179805298			OpenSea
cg15656347	chr5	7289734		chr5:7289733-7290153	Island
cg17508302	chr1	161575716	*HSPA7*	chr1:161575534-161576908	Island
cg24704510	chr17	38182512	*MED24*		OpenSea
cg22120095	chr3	50488020	*CACNA2D2*		OpenSea
cg23768702	chr18	44557673	*TCEB3CL*	chr18:44554879-44556671	S_Shore
cg23570810	chr11	315102	*IFITM1*	chr11:315739-316539	N_Shore
cg20558320	chr2	235210313			OpenSea
cg10302285	chr4	113063260		chr4:113066470-113067379	N_Shelf
cg20368377	chr13	23499360		chr13:23500589-23501398	N_Shore
cg23463608	chr19	2607757	*GNG7*	chr19:2607725-2607967	Island
cg07218880	chr13	115046279	*UPF3A*	chr13:115046754-115048034	N_Shore
cg27276613	chr11	72750477	*FCHSD2*		OpenSea
cg00810292	chr13	75900518	*TBC1D4*		OpenSea
cg19798735	chr7	110730805	*IMMP2L*		OpenSea
cg06257110	chr16	21658497	*METTL9*		OpenSea
cg11839415	chr1	43814764	*MPL*	chr1:43814305-43815277	Island
cg19627093	chr8	1900892	*ARHGEF10*		OpenSea
cg22981104	chr8	144255663			OpenSea
cg02256315	chr11	18500649	*LDHAL6A*		OpenSea
cg26234900	chr6	32820214	*TAP1*	chr6:32820849-32822370	N_Shore
cg25211348	chr1	111769665	*CHI3L2*		OpenSea
cg01367992	chr1	160766535	*LY9*		OpenSea
cg11347531	chr1	59090352			OpenSea
cg17093855	chr17	72834959	*TMEM104*	chr17:72838812-72839751	N_Shelf
cg04340595	chr1	9789174	*CLSTN1*	chr1:9790292-9790704	N_Shore
cg09026253	chr11	313267	*IFITM1*	chr11:310728-311419	S_Shore
cg06936108	chr17	7894917		chr17:7893330-7893702	S_Shore
cg03203197	chr16	34587120		chr16:34586717-34587124	Island
cg12754571	chr1	247694271	*LOC148824*	chr1:247694035-247694501	Island
cg01083549	chr16	89168371	*ACSF3*	chr16:89167116-89167482	S_Shore
cg01132696	chr6	33048558	*HLA-DPB1*	chr6:33048416-33048814	Island
cg13440692	chr1	1186357		chr1:1189396-1190245	N_Shelf
cg25170017	chr11	64644487	*EHD1*	chr11:64645429-64646566	N_Shore
cg10859442	chr11	73716367	*UCP3*		OpenSea
cg11688093	chr13	111178359	*RAB20*		OpenSea
cg02805052	chr2	104973361			OpenSea
cg08597025	chr21	15077096		chr21:15077051-15077824	Island
cg27280688	chr21	43823809	*UBASH3A*		OpenSea
cg20055861	chr15	68055293	*MAP2K5*		OpenSea
cg07212384	chr7	99761053	*GAL3ST4*		OpenSea
cg18608055	chr19	1130866	*SBNO2*		OpenSea
cg11674355	chr2	65610261	*SPRED2*		OpenSea
cg17712928	chr7	2124974	*MAD1L1*		OpenSea
cg17394978	chr5	131824502	*IRF1*	chr5:131825842-131826935	N_Shore
cg05201185	chr6	30459139	*HLA-E*	chr6:30457369-30458175	S_Shore
cg17001566	chr1	2990490	*PRDM16*	chr1:2990030-2990718	Island
cg09847153	chr1	38461540		chr1:38461584-38461988	N_Shore
cg17461271	chr1	172717683			OpenSea
cg03639671	chr4	145430689			OpenSea
cg03982998	chr1	238024671	*LOC100130331*		OpenSea
cg16351957	chr6	71874784			OpenSea
cg15787636	chr17	80866113	*TBCD*	chr17:80863327-80863743	S_Shelf

cg05830220	chr16	87757033	*KLHDC4*	chr16:87754008-87754285	S_Shelf
cg14113203	chr9	97346942	*FBP2*		OpenSea
cg04333867	chr14	24567446	*PCK2*	chr14:24563256-24563886	S_Shelf
cg08714433	chr4	132896736		chr4:132896360-132896778	Island
cg03448915	chr16	66583078	*TK2*	chr16:66583236-66584135	N_Shore
cg05347965	chr17	55663256	*MSI2*		OpenSea
cg08450017	chr3	45984838	*CXCR6*		OpenSea
cg12355172	chr11	67418315	*ACY3*		OpenSea
cg24445388	chr1	2084391	*PRKCZ*	chr1:2082314-2082529	S_Shore
cg06208288	chr1	58858074			OpenSea
cg25637655	chr6	29911542	*HLA-A*	chr6:29910202-29911367	S_Shore
cg02891314	chr5	179741120	*GFPT2*	chr5:179740710-179741121	Island
cg12048225	chr6	32808669	*PSMB8*	chr6:32811494-32811839	N_Shelf
cg05799811	chr1	167487396	*CD247*		OpenSea
cg21916358	chr21	15077562		chr21:15077051-15077824	Island
cg07318204	chr4	145566441	*HHIP*	chr4:145566242-145567413	Island
cg16727006	chr16	87470545	*ZCCHC14*		OpenSea
cg26768584	chr3	18480242	*SATB1*		OpenSea
cg13862711	chr9	124989915	*LHX6*	chr9:124987743-124991086	Island
cg26313511	chr3	125053815	*ZNF148*		OpenSea
cg09362796	chr4	3204792	*HTT*		OpenSea
cg01312482	chr5	178451176	*ZNF879*	chr5:178450630-178451216	Island
cg02932314	chr17	42733698	*C17orf104*	chr17:42733486-42734626	Island
cg03890037	chr5	115297989	*LVRN*	chr5:115297607-115298093	Island
cg04293259	chr6	30434208		chr6:30434029-30434422	Island
cg09257796	chr15	27215971	*GABRG3*	chr15:27215951-27216856	Island
cg21429394	chr12	100750899	*SLC17A8*		OpenSea
cg08513253	chr6	107955362	*SOBP*	chr6:107955163-107956693	Island
cg12277666	chr1	179560870	*TDRD5*	chr1:179560751-179561596	Island
cg12076931	chr9	139427171	*NOTCH1*	chr9:139428318-139428527	N_Shore
cg01289544	chr2	27887409	*SUPT7L*	chr2:27886259-27887410	Island
cg03896611	chr8	99986344		chr8:99985733-99986983	Island
cg16621833	chr6	117587588	*VGLL2*	chr6:117586482-117587156	S_Shore
cg20185461	chr19	18899082	*COMP*	chr19:18899037-18902284	Island
cg20692569	chr7	72848481	*FZD9*	chr7:72847933-72850032	Island
cg04626491	chr6	28584053		chr6:28583934-28584289	Island
cg23193759	chr10	71389896	*C10orf35*	chr10:71389862-71390355	Island
cg11807105	chr6	28784852		chr6:28781857-28782302	S_Shelf
cg01889893	chr1	8085193	*ERRFI1*	chr1:8085554-8086854	N_Shore
cg07906855	chr3	9642468		chr3:9642387-9642695	Island
cg03308652	chr3	62354991		chr3:62354291-62355012	Island
cg07465864	chr5	112849162	*YTHDC2*	chr5:112849265-112849982	N_Shore
cg18457731	chr8	139734484	*COL22A1*		OpenSea
cg23630442	chr8	18245225		chr8:18244455-18245226	Island
cg00986580	chr14	22951241			OpenSea
cg12190994	chr7	4305061	*SDK1*	chr7:4303079-4305062	Island
cg09862733	chr19	49631106	*PPFIA3*	chr19:49631105-49631494	Island
cg13932794	chr14	100126196	*HHIPL1*	chr14:100125671-100126694	Island
cg16178091	chr5	37844401		chr5:37836747-37840726	S_Shelf
cg13718539	chr5	140801354	*PCDHGA4*	chr5:140802399-140802832	N_Shore
cg07025583	chr5	157078881	*SOX30*	chr5:157078327-157079564	Island
cg24411302	chr14	74893183	*TMEM90A*	chr14:74892466-74893184	Island
cg08602008	chr19	48076841		chr19:48076461-48076877	Island
cg14064148	chr19	10527576		chr19:10526982-10527755	Island
cg21177456	chr19	519611	*C19orf20*	chr19:518847-519905	Island
cg07852825	chr3	172166182	*GHSR*	chr3:172165372-172166738	Island
cg15962375	chr1	213563639		chr1:213562739-213562958	S_Shore
cg20810675	chr4	171604188			OpenSea
cg06358300	chr9	84302344	*TLE1*	chr9:84302139-84305001	Island

cg24085707	chr17	79615652	*TSPAN10*	chr17:79614850-79615559	S_Shore
cg19137818	chr3	45837556	*SLC6A20*	chr3:45837480-45838256	Island
cg10304824	chr2	177028804	*HOXD3*	chr2:177029413-177029941	N_Shore
cg12050497	chr2	14773274	*FAM84A*	chr2:14772377-14775809	Island
cg08623383	chr1	28205926	*C1orf38*		OpenSea
cg15009294	chr1	214813712	*CENPF*		OpenSea
cg04096767	chr11	32449450	*WT1*	chr11:32448261-32449744	Island
cg04290586	chr12	108168987	*ASCL4*	chr12:108168986-108169570	Island
cg25782229	chr11	32450692	*WT1*	chr11:32452144-32452708	N_Shore
cg23553442	chr9	74062096		chr9:74061513-74062097	Island
cg19027255	chr10	78943401	*KCNMA1*		OpenSea
cg15726426	chr1	111216839	*KCNA3*	chr1:111216244-111217937	Island
cg10640072	chr8	99985888		chr8:99985733-99986983	Island
cg19374779	chr3	138341122	*FAIM*		OpenSea
cg15704225	chr2	81427138			OpenSea
cg22835435	chr19	51111185		chr19:51111184-51111519	Island
cg23713520	chr16	57318087	*PLLP*	chr16:57317967-57318964	Island
cg05251389	chr22	43525330	*BIK*		OpenSea
cg12289045	chr19	51141361	*SYT3*	chr19:51142284-51142986	N_Shore
cg07472704	chr13	113763758	*F7*	chr13:113761566-113765534	Island
cg20992319	chr2	225813194	*DOCK10*		OpenSea
cg14829814	chr12	108294894		chr12:108297426-108297743	N_Shelf
cg18105725	chr6	160554003	*SLC22A1*	chr6:160554661-160555504	N_Shore
cg07508429	chr7	44448689	*NUDCD3*		OpenSea
cg10424892	chr12	9602245	*DDX12*	chr12:9600415-9601259	S_Shore
cg01353448	chr7	31726912	*C7orf16*		OpenSea
cg14905403	chr17	39317349	*KRTAP4-4*		OpenSea
cg16119505	chr1	97186464	*PTBP2*	chr1:97187102-97187606	N_Shore
cg14299369	chr15	43560165	*TGM5*		OpenSea
cg16006141	chr1	25167475	*CLIC4*		OpenSea
cg10221391	chr6	29945726	*HCG9*	chr6:29944402-29945169	S_Shore
cg15115365	chr5	88384656			OpenSea
cg10207609	chr7	80267619	*CD36*		OpenSea
cg22288927	chr4	144482688		chr4:144480541-144480947	S_Shore
cg27391117	chr11	89442829	*TRIM77*		OpenSea
cg26825569	chr6	46704077	*PLA2G7*	chr6:46702736-46703316	S_Shore
cg21424209	chr6	148792600	*SASH1*		OpenSea
cg05753589	chr1	2064765	*PRKCZ*	chr1:2064628-2064855	Island
cg15136953	chr3	110793466	*PVRL3*	chr3:110790149-110791401	S_Shelf
cg18222192	chr11	79104769	*ODZ4*		OpenSea
cg00919055	chr5	149546946	*CDX1*	chr5:149546027-149546988	Island
cg14625770	chr9	104295276	*RNF20*	chr9:104295917-104296232	N_Shore
cg22109694	chr15	26360852			OpenSea
cg13629358	chr5	160112456	*ATP10B*		OpenSea
cg10992558	chr1	95369272	*CNN3*		OpenSea
cg14081737	chr9	116111377	*BSPRY*	chr9:116111663-116112189	N_Shore

Table S6- Prediction scores generated for methylation data from 60 samples with different blood cell types

Blood cell	Kabuki	ATRX	Sotos	CHARGE	Floating Harbor	ADCA-DN	Claes-Jensen	Non 7-disorder
WB_1	0.002	0.002	0.002	0.002	0.003	0.001	0.002	0.987
WB_2	0.002	0.003	0.002	0.006	0.002	0.001	0.003	0.980
WB_3	0.003	0.004	0.002	0.002	0.002	0.001	0.003	0.983
WB_4	0.002	0.002	0.002	0.001	0.002	0.001	0.002	0.990
WB_5	0.002	0.002	0.002	0.005	0.003	0.001	0.004	0.982
WB_6	0.003	0.003	0.003	0.019	0.002	0.001	0.003	0.967
PBMC_1	0.002	0.002	0.001	0.002	0.004	0.001	0.002	0.987
PBMC_2	0.001	0.002	0.002	0.002	0.002	0.001	0.003	0.987
PBMC_3	0.002	0.001	0.001	0.001	0.002	0.000	0.001	0.991
PBMC_4	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.991
PBMC_5	0.001	0.001	0.001	0.002	0.002	0.001	0.002	0.990
PBMC_6	0.002	0.001	0.001	0.005	0.002	0.001	0.001	0.986
Gran_1	0.001	0.003	0.003	0.005	0.001	0.001	0.003	0.983
Gran_2	0.003	0.003	0.004	0.030	0.003	0.001	0.007	0.950
Gran_3	0.003	0.004	0.003	0.007	0.002	0.001	0.006	0.974
Gran_4	0.002	0.003	0.003	0.002	0.002	0.001	0.004	0.982
Gran_5	0.002	0.002	0.002	0.017	0.002	0.001	0.005	0.968
Gran_6	0.006	0.006	0.004	0.078	0.003	0.002	0.006	0.895
CD4+_1	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.992
CD4+_2	0.001	0.001	0.001	0.001	0.002	0.000	0.001	0.994
CD4+_3	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.995
CD4+_4	0.001	0.001	0.001	0.001	0.002	0.000	0.001	0.994
CD4+_5	0.001	0.001	0.000	0.001	0.003	0.001	0.001	0.993
CD4+_6	0.002	0.001	0.001	0.001	0.002	0.001	0.001	0.992
CD8+_1	0.010	0.002	0.001	0.003	0.018	0.002	0.002	0.962
CD8+_2	0.003	0.001	0.001	0.001	0.005	0.001	0.001	0.988
CD8+_3	0.003	0.001	0.001	0.001	0.004	0.001	0.001	0.988
CD8+_4	0.002	0.001	0.001	0.001	0.004	0.001	0.001	0.990
CD8+_5	0.003	0.001	0.000	0.001	0.005	0.001	0.001	0.989
CD8+_6	0.006	0.001	0.001	0.002	0.004	0.001	0.001	0.984
CD14+_1	0.001	0.004	0.003	0.007	0.002	0.001	0.003	0.979
CD14+_2	0.003	0.005	0.006	0.053	0.003	0.001	0.008	0.921
CD14+_3	0.002	0.004	0.003	0.008	0.002	0.001	0.004	0.976
CD14+_4	0.002	0.002	0.003	0.002	0.001	0.001	0.003	0.987
CD14+_5	0.002	0.002	0.002	0.015	0.002	0.001	0.005	0.971
CD14+_6	0.005	0.007	0.006	0.116	0.003	0.002	0.007	0.855
CD19+_1	0.006	0.008	0.006	0.012	0.010	0.002	0.004	0.951
CD19+_2	0.004	0.006	0.008	0.025	0.006	0.001	0.004	0.944
CD19+_3	0.027	0.009	0.009	0.015	0.016	0.002	0.007	0.916

CD19+_4	0.002	0.004	0.004	0.004	0.003	0.001	0.001	0.980
CD19+_5	0.005	0.004	0.005	0.026	0.009	0.001	0.003	0.947
CD19+_6	0.013	0.013	0.016	0.251	0.009	0.003	0.006	0.688
CD56+_1	0.004	0.002	0.002	0.002	0.010	0.001	0.002	0.978
CD56+_2	0.003	0.001	0.001	0.002	0.005	0.001	0.003	0.984
CD56+_3	0.004	0.001	0.001	0.002	0.006	0.001	0.002	0.983
CD56+_4	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.992
CD56+_5	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.992
CD56+_6	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.993
Neu_1	0.002	0.004	0.003	0.006	0.003	0.001	0.005	0.977
Neu_2	0.002	0.003	0.003	0.007	0.002	0.001	0.006	0.976
Neu_3	0.003	0.004	0.003	0.005	0.002	0.001	0.007	0.976
Neu_4	0.002	0.003	0.002	0.004	0.001	0.001	0.005	0.981
Neu_5	0.002	0.002	0.002	0.008	0.002	0.001	0.004	0.981
Neu_6	0.004	0.003	0.003	0.024	0.002	0.001	0.004	0.959
Eos_1	0.001	0.004	0.003	0.004	0.002	0.001	0.004	0.981
Eos_2	0.003	0.004	0.005	0.023	0.003	0.001	0.010	0.951
Eos_3	0.002	0.004	0.004	0.006	0.002	0.001	0.008	0.974
Eos_4	0.002	0.002	0.003	0.002	0.002	0.001	0.005	0.984
Eos_5	0.002	0.002	0.002	0.007	0.002	0.001	0.006	0.978
Eos_6	0.004	0.005	0.005	0.058	0.003	0.001	0.008	0.917

WB: Whole blood, PBMC: Peripheral blood mononuclear cells, Gran: Granulocytes, Neu: Neutrophil, Eos: Eosinophil

Table S7 - Probability scores and final predictions for undiagnosed cases and VUS variants

Id	sex	age	Suspected phenotype	Screened Gene	Mutation	ACMG Variant classification	DNA methylation prediction	Kabuki	ATRX	Sotos	CHARGE	Floating Harbor	ADCA-DN	Claes-Jensen	Non 7-disorder
MS1050	m	4.3	ATRX	*ATRX*	c.2254A>G; p.T752A	VUS		0	0	0	0	0	0	0	0.99
MS0567	m	11.5	ATRX	*ATRX*	c.5579A>G; p.N1860S	Benign		0.01	0.01	0	0	0	0	0	0.97
MS1020	f	2.5	CHARGE	*CHD7*	c.5041+5G>A	VUS	CHARGE	0.03	0.01	0.01	0.82	0.02	0.01	0.01	0.1
MS0978	f	0.8	CHARGE	*CHD7*	c.2788G>A, p.E930K	VUS		0	0	0	0	0.01	0	0	0.98
MS0980	m	16	CHARGE	*CHD7*	c.6775+1G>A	VUS	CHARGE	0.01	0.01	0.01	0.94	0.01	0.01	0.01	0
MS0981	m	3.5	CHARGE	*CHD7*	c.2058insGCAAAA	VUS		0.02	0	0	0.01	0	0	0	0.97
MS1028	m	4.3	CHARGE	*CHD7*	c.2185A>G, p.K729E	VUS	Kabuki	0.92	0.01	0.01	0.01	0.02	0.01	0.01	0.02
MS1031	f	7.5	CHARGE	*CHD7*	c.5405-17G>A (de novo)	Pathogenic	CHARGE	0.04	0.02	0.01	0.65	0.01	0.01	0.01	0.25
MS1034	m	0.01	CHARGE	*CHD7*	c.5405-17G>A	Pathogenic	CHARGE	0.02	0.02	0.01	0.89	0.01	0.01	0.01	0.02
MS0986	m	0.01	CHARGE	*CHD7*	c.5405-17G>A	VUS		0.02	0.04	0.01	0.2	0.01	0.01	0	0.71
MS0988	f	2.5	CHARGE	*CHD7*	c.3990-58insA, c.5210+58insA, c.5535-68T>C	VUS		0.01	0	0	0.07	0	0	0	0.91
MS0990	m	10.1	CHARGE	*CHD7*	c.6104-22_6104-4del19	VUS	CHARGE	0.03	0.03	0.02	0.82	0.03	0.01	0.03	0.03
MS0993	f	1.5	CHARGE	*CHD7*	c.2957+5delG	VUS		0.02	0.01	0	0.17	0.01	0.01	0	0.79
MS1042	f	0.9	CHARGE	*CHD7*	c.2124T>C, p.S708S	VUS		0.01	0.01	0	0.15	0.01	0	0	0.81
MS1043	m	0.2	CHARGE	*CHD7*	c.1565G>T, pG522V	VUS		0	0	0	0.07	0	0	0	0.92
MS1047	m	3	CHARGE	*CHD7*	No mutation was found in CHD7	N/A	CHARGE	0.02	0.01	0.01	0.91	0.02	0.01	0.01	0.02
MS0994	f	13.1	CHARGE	*CHD7*	c.712G>A, p.V238M	VUS		0.04	0.02	0	0.01	0.01	0.01	0	0.91
MS0997	f	0.4	CHARGE	*CHD7*	c.5534+16T>C	VUS		0	0.01	0	0	0	0	0	0.97
MS0998	f	4.4	CHARGE	*CHD7*	c.2915A>G, Q972R	VUS	CHARGE	0.04	0.02	0.01	0.82	0.03	0.01	0.02	0.05
MS0999	f	0.3	CHARGE	*CHD7*	c.2053_2058dup6, c.5517dupT	VUS	CHARGE	0.02	0.01	0.01	0.92	0.01	0.01	0.02	0.01
MS1000	f	0.01	CHARGE	*CHD7*	c.2490C>A	VUS	CHARGE	0.02	0.02	0.01	0.92	0.01	0.01	0.01	0.01
MS1001	m	0.2	CHARGE	*CHD7*	c.4533+1G>A	Pathogenic	CHARGE	0.03	0.02	0.01	0.8	0.01	0.01	0.03	0.09
MS1006	m	0.001	CHARGE	*CHD7*	c.6321C>T, p.H2107H	VUS		0.01	0.01	0.01	0.18	0.01	0.01	0	0.78
MS1008	f	0.9	CHARGE	*CHD7*	c.6937-1G>A	Pathogenic	CHARGE	0.04	0.04	0.03	0.82	0.02	0.02	0.02	0.02
CHD7-20	f	5	CHARGE	*CHD7*	c.6322G>T (p.Gly2108Trp)	VUS		0.02	0.01	0	0.15	0.01	0	0	0.75

CHD7-21	f	3	CHARGE	*CHD7*	c.3746G>A (p.Arg1249Gln)	VUS	CHARGE	0.02	0.01	0.01	0.73	0.02	0	0.01	0.2
CHD7-22	f	0.01	CHARGE	*CHD7*	c2751G>A (p.Thr917Thr)	VUS		0	0	0	0.08	0	0	0	0.9
CHD7-23	m	0.5	CHARGE	*CHD7*	c.-15G>A	VUS	CHARGE	0.02	0.01	0.01	0.69	0.01	0.01	0.01	0.25
CHD7-24	m	2	CHARGE	*CHD7*	c.4225G>A (p.Val1409Met)	VUS	CHARGE	0.02	0.01	0.01	0.74	0.01	0	0.01	0.2
CHD7-25	m	13	CHARGE	*CHD7*	c.5436C>G (p.Asp1812Glu)	VUS	CHARGE	0.02	0.02	0.02	0.83	0.01	0.01	0.02	0.07
CHD7-26	m	4	CHARGE	*CHD7*	c.5633A>G (p.Asp1878Gly)	VUS		0	0	0	0	0	0	0	0.99
CHD7-27	m	4	CHARGE	*CHD7*	c.5848G>A (p.Ala1950Thr)	VUS		0.01	0	0	0.02	0	0	0	0.97
CHD7-28	m	13	CHARGE	*CHD7*	c.6304G>T (p.Val2102Phe)	VUS		0	0	0	0.01	0	0	0	0.99
CHD7-29	m	6	CHARGE	*CHD7*	c.3566G>A (p.Arg1189His)	VUS		0	0	0	0.03	0	0	0	0.95
CHD7-30	m	0.08	CHARGE	*CHD7*	intron4:c2238+1del	VUS	CHARGE	0.01	0.01	0.01	0.95	0	0	0.01	0
CHD7-31	m	5	CHARGE	*CHD7*	c.2049_2050insAAAGCA (p.Ala685_Lys686dup)	VUS		0	0	0	0	0	0	0	0.99
CHD7-32	m	9	CHARGE	*CHD7*	c.6377A>T (p.Asp2126Val)	VUS		0	0	0	0.04	0	0	0	0.95
CHD7-33	m		CHARGE	*CHD7*	c.4851T>G (p.=(p.Gly1617Gly))	VUS		0	0	0	0.02	0	0	0	0.97
CHD7-35	f		CHARGE	*CHD7*	c.8791 G>A (p.Val2931Met)	VUS	CHARGE	0.02	0.01	0.01	0.86	0.01	0	0	0.09
CHD7-37	m		CHARGE	*CHD7*	c.8802C>G (p.Ser2934Arg)	VUS		0	0	0	0.02	0	0	0	0.98
CHD7-38	m		CHARGE	*CHD7*	c.2516_2518delAGT (p.Gln839_Trp840delinsArg)	VUS	CHARGE	0.03	0.02	0.01	0.77	0.02	0.01	0.01	0.13
CHD7-39	f		CHARGE	*CHD7*	c.8759G>C (p.Gly2920Ala)	VUS		0	0	0	0	0	0	0	0.99
CHD7-42	f		CHARGE	*CHD7*	c.127A>G (p.Ile43Val)	VUS	CHARGE	0.01	0.01	0	0.96	0.01	0	0	0
CHD7-43	m		CHARGE	*CHD7*	c.1405A>G (p.Arg469Gly)	VUS		0	0	0	0	0	0	0	0.99
CHD7-44	f		CHARGE	*CHD7*	c.7763A>G (p.Asn2588Ser)	VUS	CHARGE	0.01	0.01	0	0.97	0	0	0.01	0
CHD7-45	m		CHARGE	*CHD7*	c.3871A>C (p.Lys1291Gln)	VUS	CHARGE	0.01	0.01	0.01	0.95	0.01	0	0.01	0
CHD7-48	m		CHARGE	*CHD7*	c.6193C>T (p.Arg2065Cys)	VUS	CHARGE	0.02	0.02	0.01	0.89	0.01	0.01	0.01	0.04
CHD7-49	f		CHARGE	*CHD7*	c.3762T>A (p.His1254Gln)	VUS	CHARGE	0.02	0.01	0.01	0.91	0.02	0	0.01	0.01
CHD7-50	m		CHARGE	*CHD7*	c.583C>T (p.Arg195Cys)	VUS		0	0	0	0.01	0	0	0	0.99
CHD7-54	f		CHARGE	*CHD7*	c.1562C>T (p.Pro521Leu)	VUS		0	0	0	0.03	0	0	0	0.97
CHD7-57	f		CHARGE	*CHD7*	c.5666-9C>G	VUS	CHARGE	0.01	0.01	0.01	0.96	0.01	0	0.01	0
CHD7-59	m		CHARGE	*CHD7*	c.1797_1799delGAA (p.Lys602del)	VUS		0	0	0	0	0	0	0	0.99
CHD7-62	f		CHARGE	*CHD7*	c.5429G>C (p.Arg1810Pro)	VUS	CHARGE	0.03	0.01	0	0.77	0.01	0.01	0.01	0.16
CHD7-65	m		CHARGE	*CHD7*	c.5827C>T (p.Arg1943Trp)	VUS		0.01	0	0	0.08	0	0	0.01	0.89

CHD7-66	m		CHARGE	*CHD7*	c.317A>G (p.His106Arg)	VUS		0.01	0	0	0.11	0	0	0	0.86
CHD7-67	m		CHARGE	*CHD7*	c.6529G>A (p.Gln2177Lys)	VUS		0.01	0.02	0.02	0.15	0.01	0	0.01	0.79
CHD7-70	f		CHARGE	*CHD7*	c.6356A>G (p.Asp2119Gly)	VUS	CHARGE	0.03	0.01	0.01	0.81	0.02	0.01	0.01	0.1
MS0694	f	54	Claes-Jensen - carrier	*KDM5C*	c.1510G>A; p.V504M	Pathogenic		0.03	0.01	0.01	0.03	0.02	0.01	0.1	0.8
MS0696	f	31	Claes-Jensen - carrier	*KDM5C*	c.1510G>A; p.V504M	Pathogenic		0.01	0.01	0	0.01	0.01	0	0.04	0.92
MS0725	f	66	Claes-Jensen - carrier	*KDM5C*	c.229G>A; p.A77T	Pathogenic		0.03	0.01	0.01	0.03	0.01	0.01	0.11	0.79
MS0729	f	17	Claes-Jensen - carrier	*KDM5C*	c.229G>A; p.A77T	Pathogenic		0.02	0.01	0.01	0.03	0.02	0	0.1	0.82
MS0730	f	51	Claes-Jensen - carrier	*KDM5C*	c.4439_4440DdelAG; p.fsR1481fxX9	Pathogenic		0.02	0.01	0	0.02	0.01	0	0.05	0.89
MS0731	f	39	Claes-Jensen - carrier	*KDM5C*	c.229G>A; p.A77T	Pathogenic		0.06	0.03	0.02	0.1	0.04	0.02	0.27	0.66
MS0732	f	55	Claes-Jensen - carrier	*KDM5C*	c.229G>A; p.A77T	Pathogenic		0.02	0.01	0.01	0.03	0.01	0	0.08	0.85
MS0733	f	54	Claes-Jensen - carrier	*KDM5C*	c.1510G>A; p.V504M	Pathogenic		0.05	0.02	0.01	0.06	0.03	0.01	0.14	0.68
KDM6A	f		Kabuki	*KDM6A*	c.2668_2669dupTA (p.Pro891Thrfs*8)	Pathogenic	Kabuki	0.65	0.01	0.01	0.07	0.02	0.01	0.02	0.21
MS0710	f	9.1	Kabuki	*KMT2D*	c.8382C>A (p.G2794G)	Likely_Benign		0	0	0	0	0	0	0	0.99
MS0712	m	10.8	Kabuki	*KMT2D*	c.2438C>T (p.P813L)	Benign		0	0	0	0	0	0	0	0.99
MS0714	f	14.4	Kabuki	*KMT2D*	c.15540G>C (p.V5180V)	Likely_Benign		0	0	0	0	0	0	0	0.99
MS0715	f	9.3	Kabuki	*KMT2D*	c.5868-25G>A	Likely_Benign		0.24	0.01	0	0.02	0.01	0	0.01	0.71
MS0716	f	1.3	Kabuki	*KMT2D*	c.16229G>A (p.G5410E)	VUS	Kabuki	0.9	0.01	0.01	0.02	0.01	0.01	0.01	0.04
MS0717	f	9.2	Kabuki	*KMT2D*	c.1938C>G (p.P646P)	Likely_Benign		0.07	0	0	0.01	0	0	0	0.91
MS0718	m	2.8	Kabuki	*KMT2D*	c.4265G>C (p.W1422S)	VUS	Kabuki	0.62	0.02	0.01	0.03	0.02	0.01	0.01	0.3
MS0734	f	3.6	Kabuki	*KMT2D*	c.15641G>T (p.R5214L)	VUS	Kabuki	0.9	0.01	0.01	0.01	0.02	0.01	0.01	0.04
MS0736	f	1.4	Kabuki	*KMT2D*	c.3907-15C>G	Likely_Benign		0.06	0.01	0	0.01	0.01	0	0.01	0.9
MS0742	m	1.2	Kabuki	*KMT2D*	c.11750_11758del9	VUS		0	0	0	0	0	0	0	0.98
MS0791	f	2.8	Kabuki	*KMT2D*	c.10741-42G>A	Likely_Benign		0	0	0	0	0	0	0	0.99
MS0797	m	13.8	Kabuki	*KMT2D*	c.12889T>C (p.S4297P)	VUS		0	0	0	0	0	0	0	0.99
MS0775	m	1.8	Kabuki	*KMT2D*	c.4143G>A (p.V1381V)	VUS		0	0	0	0	0	0	0	0.99
MS0777	m	13.4	Kabuki	*KMT2D*	c.401-11delC	Likely_Benign		0.02	0	0	0.01	0	0	0	0.97
MS0779	m	6.3	Kabuki	*KMT2D*	c.12913G>A (p.V4305I)	Likely_Benign		0	0	0	0	0	0	0	0.99
MS0780	m	1.3	Kabuki	*KMT2D*	c.6752C>T (p.S2251L)	Likely_Benign		0.01	0	0	0	0	0	0	0.99

MS0781	f	4.3	Kabuki	*KMT2D*	c.15108_15110del3 (p.H5036_E5037delinsQ)	VUS	Kabuki	0.94	0.01	0	0.01	0.01	0.01	0.01	0.01
MS0785	m	7.7	Kabuki	*KMT2D*	c.13644C>T (p.S4548S)	Benign		0	0	0	0.01	0	0	0	0.98
KMT2D-12	f	4.5	Kabuki	*KMT2D*	c.15143G>A (p.Arg5048His)	VUS	Kabuki	0.9	0.01	0.01	0.02	0.01	0.01	0.01	0.04
KMT2D-13	m	5	Kabuki	*KMT2D*	c.12028 T>C (p.Ser4010Pro)	VUS		0	0	0	0.01	0	0	0	0.98
KMT2D-14	f	15	Kabuki	*KMT2D*	c.16522-5_16522-4delTT	VUS	Kabuki	0.69	0.01	0.01	0.06	0.01	0.01	0.02	0.19
KMT2D-15	m	14	Kabuki	*KMT2D*	c.15910A>G (p.Ile5304Val)	VUS		0.01	0	0	0.17	0.01	0	0.01	0.79
KMT2D-16	m	17	Kabuki	*KMT2D*	c.15659G>A (p.Arg5220His)	VUS		0	0	0	0.02	0	0	0	0.96
KMT2D-17	m	9	Kabuki	*KMT2D*	c.10256A>G (p.Asp3419Gly)	VUS		0	0	0	0.01	0	0	0	0.98
KMT2D-18	f	16	Kabuki	*KMT2D*	c.8974G>A (p.Glu2992Lys)	VUS		0	0	0	0	0	0	0	0.99
KMT2D-19	f	8	Kabuki	*KMT2D*	c.8831A>G (p.Asn2944Ser)	VUS		0	0	0	0.01	0	0	0	0.98
KMT2D-20	f	10	Kabuki	*KMT2D*	c.832G>A (p.Ala278Thr)	VUS		0	0	0	0.01	0	0	0	0.99
KMT2D-21	m	6	Kabuki	*KMT2D*	c.682C>G (p.Arg228Gly)	VUS		0	0	0	0.01	0	0	0	0.98
KMT2D-23	m		Kabuki	*KMT2D*	c.15587T>A (p.Met5196Lys)	VUS		0	0	0.01	0.03	0	0	0	0.95
KMT2D-24	f		Kabuki	*KMT2D*	c.11150A>C (p.Gln3717Pro)	VUS		0	0	0	0	0	0	0	0.98
KMT2D-27	m		Kabuki	*KMT2D*	c.11578_11580dupCAG (p.Gln3863dup)	VUS		0	0	0	0.06	0	0	0	0.93
KMT2D-28	m		Kabuki	*KMT2D*	c.10909C>A (p.Pro3637Thr)	VUS		0.01	0.01	0	0.22	0	0	0	0.75
KMT2D-31	m		Kabuki	*KMT2D*	c.12662_12664dupAGC (p.Gln4221dup)	VUS		0	0	0	0.03	0	0	0	0.96
KMT2D-34	f		Kabuki	*KMT2D*	c.2334C>G (p.Cys778Trp)	VUS		0	0	0	0.1	0	0	0	0.88
KMT2D-35	f		Kabuki	*KMT2D*	c.14731_14733delCCT (p.Pro4911del)	VUS		0	0	0	0	0	0	0	0.99
KMT2D-37	f		Kabuki	*KMT2D*	c.15626G>T (p.Gly5209Val)	VUS	Kabuki	0.79	0.01	0.01	0.04	0.01	0.01	0.01	0.11
DL141677	m	18	Sotos_fibroblast	*NSD1*	ex20-21 del (1860-2086 aa)	Pathogenic	Sotos	0.03	0.1	0.65	0.03	0.05	0.01	0.03	0.1
51537F	f	1.6	Sotos_fibroblast	*NSD1*	c.1810C>T (p.Arg604*)	Pathogenic	Sotos	0.05	0.1	0.58	0.04	0.05	0.01	0.03	0.13
SF155	f		Sotos_fibroblast	*NSD1*	c.2362C>T (p.Arg788*)	Pathogenic	Sotos	0.06	0.1	0.62	0.04	0.04	0.01	0.03	0.1
DL208122	m	2.3	Sotos	*NSD1*	c.6437G>C (p.Cys2146Ser)	VUS	Sotos	0.03	0.03	0.68	0.06	0.01	0.01	0.02	0.17
DL136303	f	5	Sotos	*NSD1*	c.5779G>C (p.Ala1927Pro)	VUS	Sotos	0.01	0.02	0.86	0.04	0.01	0.01	0.01	0.04
DL199861	m	7	Sotos	*NSD1*	c.6412T>C (p.Cys2138Arg)	VUS	Sotos	0.03	0.03	0.65	0.06	0.02	0.01	0.02	0.19
DL181344	m	13	Sotos	*NSD1*	c.4949A>G (p.Asn1650Ser)	VUS		0	0	0	0	0	0	0	0.99
DL73286	m	10	Sotos	*NSD1*	c.3722G>C (p.Ser1241Thr)	VUS		0	0	0	0.01	0	0	0	0.99

DL159025	m	2.5	Sotos	*NSD1*	c.6013C>G (p.Arg2005Gly)	VUS	Sotos	0.01	0.02	0.9	0.02	0.01	0.01	0.01	0.03
HK-9776	f		Sotos	*NSD1*	c.1070A>G (p.Asn357Ser)	VUS		0.01	0	0	0.01	0	0	0	0.98
HK-12366	f		Sotos	*NSD1*	c.3446A>G (p.Asn1149Ser)	VUS		0	0	0	0.02	0	0	0	0.97
HK-5474	f		Sotos	*NSD1*	c.4817G>A (p.Cys1606Tyr)	VUS	Sotos	0.01	0.02	0.84	0.06	0.01	0.01	0.01	0.03
HK-11693	f		Sotos	*NSD1*	c.5177C>G (p.Pro1726Arg)	VUS	Sotos	0.01	0.02	0.9	0.02	0.01	0	0.01	0.04
HK-11767	f		Sotos	*NSD1*	c.5357A>G (p.Lys1786Arg)	VUS		0	0	0	0	0	0	0	0.99
HK-5581	m		Sotos	*NSD1*	c.5903T>C (p.Val1968Ala)	VUS		0.03	0.02	0.1	0.17	0.01	0.01	0.02	0.59
HK-3326	f		Sotos	*NSD1*	c.5990A>G (p.Tyr1997Cys)	VUS	Sotos	0.02	0.02	0.81	0.04	0.01	0.01	0.01	0.08
HK-435	m		Sotos	*NSD1*	c.6049C>T (p.Arg2017Trp)	VUS	Sotos	0.01	0.02	0.89	0.03	0.01	0.01	0.01	0.03
HK-6943	m		Sotos	*NSD1*	c.6674C>A (p.Pro2225Gln)	VUS		0	0	0	0.01	0	0	0	0.99
HK-14867	m		Sotos	*NSD1*	c.7421A>G (p.Gln2474Arg)	VUS		0	0	0	0	0	0	0	0.99
MS0436	f	40	Floating Harbor (Sex mismatch)	*SRCAP*	p.Arg2748*	Pathogenic	Floating Harbor	0.02	0.02	0.01	0.02	0.87	0.01	0.02	0.03

Reference sequence for all mutations is hg19, except for NSD1 mutations (hg18)