

The Beauty of Science is to Make Things Simple

Genome-Wide DNA Methylation Analysis In Plants and Animals

Xueguang Sun, TzuHung Chung, Eliza Bacon, Ron Leavitt, Nikolas Isely, Marc Van Eden & Xi-Yu Jia

Zymo Research Corporation, Irvine, CA

Abstract

DNA methylation is a highly conserved epigenetic mark present in many eukaryotic organisms including plants, animals, and It plays an important role in the regulation of gene fungi. expression. A number of studies have shown its involvement in plant and animal growth and reproduction primary through the processes of genomic imprinting, X-chromsome inactivation, and the silencing of transposons and other repetitive DNA elements. As such, the understanding of DNA methylation has become a major focus of the research conducted during the "post-genomics" era. However, the precise determination of a DNA's methylation pattern on a genomic scale has posed a challenge especially for those complex genomes. Combining well-established bisulfite conversion chemistry with NextGen sequencing, we have established a robust service platform for analyzing DNA methylation with single base resolution at the genomic scale. The service features a streamlined workflow coupled with a comprehensive bioinformatics pipeline to provide both a consolidated and costeffective solution for epigenetic analysis of plant and animal This technology has been used successfully for the genomes. analysis of methylomes from many organisms including soybean, mouse, and chicken. Additional analyses of other species are ongoing. Those data should provide a means to understanding how environment as well as other factors may alter an organism's fitness through heritable changes in epigenetic gene expression.

Service comparison

EPIQUEST NEXT-GE	N SEQUENCING DNA METHYLATION ANALYSIS	BASIC S	SERVICE (AGE	FULL SERVICE PACKAGE				
Service Performed	vice Performed Description							
Genomic DNA Sample Validation*	Quality validation of samples supplied by client prior to processing.	~	~	✓	~			
Sample Processing & Library Preparation	Unique one-step procedure for optimized performance and limited sample loss: endonuclease digestion, end modification, adaptor addition, size selection, bisufite treatment (conversion), limited amplification.	~	~	~	~			
Next-Gen Sequencing	Next-Gen sequencing using the latest technologies available from Illumina®.	1	✓	~	~			
Sequencing Validation & Bioinformatic Analysis*	Alignment of sequencing data to the corresponding reference genome using proprietary, state of the art bioinformatic algorithms and software.	~	~	~	~			
Data Output & Delivery	Delivery of a consolidated printed report including: service checklist, performance metrics, statistics, select graphical data, and more. Project data will be provided to the client on a hard drive.	~	×	~	~			
Data Provided	Description							
Genome Browser Tracks	Three (3) UCSC browser tracks: CpG track, Next-Gen sequencing read track, and endonuclease-digested fragment track. Reference UCSC Genome Browser: (http://genomes.ucsc.edu)	~	~	~	~			
Sequence Coverage Information	Statistical coverage (pie chart) of genomic features analyzed (i.e., CpG Islands, promoter regions, etc.)			~				
Hierarchical Clustering	Distribution (heat map) of methylated vs. non-methylated loci between two (2) or more samples.			~				
Methylation Frequency	Pairwise sample comparisons (barchart) of methylation frequencies in CpG islands, promoter regions, etc.			~	Inquire			
Methylation Index	Global methylation (percentage) for CpG Islands, promoters, etc.			✓				
Differentially Methylated Regions	Index of hyper- and hypo-methylated regions for biomarker candidate determination.			✓				
Correlation Data	Pearson correlation for pairwise data comparison between samples.			~				
and much more!	Our unique bioinformatics pipeline is used to generate >35 publication-ready figures and tables for easy data interpretation and sample comparison.			~				
	Inqu	uire†	Inqu	Inquire [†]				
ADDITIONAL CUSTOMIZED SERVIC	ES ARE AVAILABLE							
Non-Human & -Mouse Sampling	Samples from organisms other than human and mouse having annotated genomes can also be processed with the EpiQuest [™] Genome-Wide DNA Methylation workflow.							
DNA Extraction/Purification from Difficult Sample Sources	DNA can be isolated from cells, blood, tissue, FFPE tissue, environmental sources, and more!	Inquire						
Expanded Genomic Analyses	Analysis of gene bodies, repeated elements, exons, introns, non-coding RNAs, miRNAs, enhancers, and more!							
Analysis of Non-CpG Methylation	Expanded analysis of CHG and/or CHH methylation.							
Customized Bioinformatic	Our bioinformaticians can customize data delivery and graphics upon request.							

hyper/hypo methylation gene list

				Emb1 PGC1				Emb1 vs. PGC1							Emb1			PGC1			Emb1 vs. PGC1				
D	chrom	chromstart	chromend	Meth Ratio	Total CpGs	Unique CpGs	Meth Ratio	Total CpGs	Unique CpGs	Methylation Ratio Diff	pValue	Classification	ID	chroi	n chromsta	t chromend	Meth Ratio	Total CpGs	Unique CpGs	Meth Ratio	Total CpGs	Unique CpGs	Methylation Ratio Diff	pValue	Classification
94	chr1	79541375	79541928	0.9122332	546	5	0.0511609	640	5	0.86107	2.76E-234	stronglyHypermeth	108	1 chr1	8301552	83015724	0	57	3	1	2	2	-1	5.84E-04	stronglyHypomet
97	chr7	1880724	1880937	0.875	8	8	0.0520833	93	8	0.82292	3.51E-07	stronglyHypermeth	533	4 chr4	470954	471159	0	4	4	1	39	4	-1	8.10E-06	stronglyHypomet
492	chr10	32354	33231	0.9481061	92	11	0.1519841	92	12	0.79612	4.50E-21	stronglyHypermeth	840	4 chré	3409479	34095134	0	25	2	1	20	2	-1	3.15E-13	stronglyHypomet
502	chrZ	71798662	71799033	0.8473568	1315	5	0.0534008	1045	5	0.79396	0.00E+00	stronglyHypermeth	1431	L6 chr1	7 7722081	7722537	0.0090909	89	10	1	80	4	-0.99091	9.19E-47	stronglyHypomet
95	chr1	79542904	79543350	0.8097766	9129	32	0.0259019	15379	35	0.78387	0.00E+00	stronglyHypermeth	1931	L1 chrž	5458342	54583659	0.0123457	85	3	1	33	3	-0.98765	1.78E-28	stronglyHypomet
93	chr1	79535777	79537276	0.8379483	132	4	0.0665205	150	4	0.77143	3.17E-42	stronglyHypermeth	757	6 chr	5470543	54705708	0.0119048	88	9	0.9961538	103	10	-0.98425	1.60E-49	stronglyHypomet
682	chr10	6063757	6064070	0.8382143	77	10	0.087031	195	12	0.75118	7.18E-37	stronglyHypermeth	929	2 chr8	5960366	5960853	0.020202	107	10	0.9923077	111	10	-0.97211	1.12E-57	stronglyHypomet
496	chr13	10456519	10456785	0.7605311	122	14	0.0125	148	10	0.74803	5.05E-37	stronglyHypermeth	1119	97 chr1	1 1113554	1113958	0.0138889	284	13	0.9856366	147	8	-0.97175	1.46E-107	stronglyHypomet
243	chr13	836790	837765	0.7669627	350	23	0.0252523	672	26	0.74171	6.18E-166	stronglyHypermeth	1249	91 chr1	3 1042778	10428017	0.015873	71	9	0.9875	165	9	-0.97163	1.54E-56	stronglyHypomet
11	chr4	7087287	7087775	0.7817155	217	15	0.0617284	159	15	0.71999	1.37E-43	stronglyHypermeth	1371	L4 chr1	5 7746468	7746928	0.0164835	65	7	0.9867038	53	7	-0.97022	3.77E-27	stronglyHypomet
036	chr27	1183579	1184131	0.8	28	10	0.0868946	65	6	0.71311	7.03E-16	stronglyHypermeth	1879	94 chrž	1122937	1123202	0.021164	91	9	0.9907834	67	7	-0.96962	1.28E-32	stronglyHypomet
587	chr12	52139	52951	0.7509345	541	29	0.055267	1368	29	0.69567	1.51E-240	stronglyHypermeth	1575	51 chr2	5306581	5306957	0	12	3	0.9666667	30	3	-0.96667	1.18E-09	stronglyHypomet
498	chrZ	71697511	71698066	0.821115	710	5	0.1337576	799	5	0.68736	9.15E-152	stronglyHypermeth	513	3 chr	10293588	4 102936269	0.0242347	190	14	0.9882576	152	12	-0.96402	1.48E-85	stronglyHypomet
955	chr27	28893	29346	0.750568	537	5	0.0883531	2346	5	0.66221	1.52E-225	stronglyHypermeth	1029	97 chr	1711299	17113231	0.0268966	152	5	0.9866667	95	5	-0.95977	6.85E-61	stronglyHypomet
325	chr14	14401560	14403554	0.7305698	291	11	0.0729431	678	11	0.65763	2.09E-73	stronglyHypermeth	1030	03 chr	1713458	17134785	0.0416667	30	8	1	44	8	-0.95833	9.59E-20	stronglyHypomet
323	chr14	14383362	14385380	0.7466794	728	10	0.0907709	1524	10	0.65591	6.33E-247	stronglyHypermeth	1004	15 chr	8779396	8779830	0.0444444	42	3	1	39	3	-0.95556	4.25E-21	stronglyHypomet
99	chr7	2124488	2125129	0.6603454	140	14	0.0054945	148	13	0.65485	1.17E-14	stronglyHypermeth	142	6 chr1	11607486	7 116075120	0.0296703	30	5	0.9846154	47	5	-0.95495	1.55E-17	stronglyHypomet
86	chr5	55446106	55446429	0.7651515	47	4	0.125	5	4	0.64015	2.06E-02	stronglyHypermeth	1258	37 chr1	3 1343986	13440067	0	38	4	0.9522727	43	4	-0.95227	4.25E-21	stronglyHypomet
303	chr11	4527633	4528265	0.7057089	359	29	0.0674149	570	25	0.63829	2.34E-90	stronglyHypermeth	1874	12 chr2	4184186	4184441	0.038961	50	3	0.990991	57	3	-0.95203	8.15E-27	stronglyHypomet
393	chr12	7155773	7156222	0.7248789	91	15	0.0909909	172	11	0.63389	2.34E-18	stronglyHypermeth	102	0 chr1	8029157	8 80292161	0.0091503	221	17	0.9601393	224	17	-0.95099	1.25E-117	stronglyHypomet

Introduction

The EpiQuest^M Genome-Wide DNA methylation Profiling Service expands upon the conventional RRBS (Reduced-Representation Bisulfite Sequencing) method to greatly increase sequence analysis of CpG-riched DNA. Through an unique library preparation procedure and an optimized workflow, the EpiQuest^M team at Zymo Research has been able to expand coverage to \geq 80% of all gene promoters and \geq 85% of CpG islands (for human samples).

Genome-wide 5mC profiling in chicken

Cost (per sample



Pair-wise comparison was performed between the different samples, DMR were counted according to hyer/hypo methylation.

Whole Genome 5mC profiling in soybean



Genome coverage



Coverage per cytotine

The EpiQuest[™] Whole-Genome DNA methylation Profiling Service utilizes a streamlined library preparation and sophisticated alignment algorithms to deliver Whole Genome Bisulfite-Sequencing data that covers >80% of all the bases in the human genome at single-base resolution.





CpG island

Gene promoter

Pie charts shows the overall distribution of sequencing reads coverage

UCSC genome browser track



Methylation Index



> Distribution of CG/CHG/CHH sites for three soybean samples





Metaplot for gene, exon and CpG island





CHG $\begin{array}{c} & & & & \\$



Genome Coverage :



Next-Gen sequencing based platforms

✓ Low DNA input

- ✓ Applicable to a broad range of sample sources (human, mouse, rat, plant, etc.)
- ✓ Streamlined workflows with comprehensive bioinformatic analysis and high quality publishable data delivery.
- \checkmark Customizable, rapid turnaround at an affordable cost.





Correlation analysis



> Hierarchical clustering:



Heatmap shows the hierarchical clustering by region and sample Pearson correlation for pairwise data comparison between samples

Emb1.RRBS.methRatio



- References
- 1. Meissner A, Mikkelsen TS, Gu H, Wernig M etc. Nature. 2008 Aug 7;454(7205):766-70.
- 2. Smith ZD, Gu H, Bock C, Gnirke A, Meissner A. Methods. 2009 Jul;48(3):226-32