

SUPPLEMENTAL INFORMATION

Supplemental Contents:

Supplemental Materials and Methods

Supplemental Table 1: Sequence information for the clustering of other GATC sites around known unmethylated GATC sites

Supplemental Figure 1: Histogram summarizing the data from Supplemental Table 1

Supplemental Figure 2: Validation of DpnI assay

Supplemental Figure 3: Modeling of different ratios of k_1 and k_2 (From Figure 3D, primary manuscript) to explore the extent of intrasite processivity

Supplemental Figure 4: Sample PAGE gels

SUPPLEMENTAL MATERIALS AND METHODS

DpnI digestion control experiment: Single stranded DNA and its reverse complement were ordered and annealed (5'-ATCGTGGACTTCTACTTGGATGGAGAAAAG**GATCGAC**ACGTATTCCAGGAATTCACGTTAC-3') generating substrate H. The DNA was then subjected to M.Taq methylation. The 50 uL reaction contained 210 nM DNA, 100 µg/mL BSA, 30uM SAM, 5uL NEB Buffer 4, and 30 units of M.Taq (NEB); it was incubated for 3 hours at 65 degrees. The reaction was then subjected to a phenol chloroform extraction and then an ethanol precipitation, making substrate H_{hm}. M.Taq methylates adenines in 5'-TCGA-3' sequences (underlined), creating a hemimethylated GATC site (bold). To show that the substrate was completely hemimethylated, the DNA was subjected to DpnII (NEB) digestion. The hemimethylated DNA was then subjected to Dam methylation under single turnover reaction conditions (see primary manuscript) for several hours to insure complete methylation, making substrate H_{dm}. The H_{hm} control lane in Supplemental Figure 2B was incubated with the same single turnover conditions as H_{dm}, but no AdoMet was added. 2.5µL of each reaction was added into 14.8 µL of water. 2 µL of NEB Buffer 4 was then added. Then 0.7 µL (14 units) of DpnI (NEB) was added and the reaction was incubated at 37 degrees for 10 minutes. The reactions were heat killed and slow cooled to room temperature before PAGE (see primary manuscript).

SUPPLEMENTAL TABLE LEGENDS

Supplemental Table 1: A List of known heritably unmethylated GATC sites (underlined) and the distance to their closest adjacent GATC site. Most (19 out of 26) of these sites have been implicated to be involved in gene regulation because they overlap with binding sites for known regulatory proteins. Examples where the spacing between GATC sites is greater than 200 base pairs are not included. ^a sequence from E. coli k-12 chromosome MG155. nc 000913.2; “number” represents list from table in (1).

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure 1: Histogram of variable clustering of GATC sites (compiled from Supplemental Table 1).

Supplemental Figure 2: Validation of the DpnI cutting assay. (A) Substrate H is methylated by M.Taq to create substrate Hhm (hemimethylated GATC site on DNA). Lane 1 is substrate Hhm (hemi methylated 60mer) subjected to DpnII digestion. DpnII cuts unmethylated GATC sites. Lane 2 is substrate H subjected to DpnII digestion (digestion fragments are both ~30 base pairs and are shown overlaid on the gel). Protection from cutting shows that substrate Hhm is completely hemi methylated. (B) Substrate Hhm and substrate Hdm (completely methylated GATC site on DNA) are subjected to DpnI digestion. Lane 1 shows that substrate Hhm is uncut while substrate Hdm is significantly cut.

Supplemental Figure 3: Traces for different ratios of $k_1:k_2$ used to model the data from Figure 3D. The black line is the tritium data; the black dots are $k_2=1000k_1$ and $k_2=100k_1$ (both traces are identical); the grey dots are $k_2=10k_1$; the grey diamonds are $k_2=.1k_1$; the grey dashed line is sequential methylation, $k_1 \sim k_2$. Notably, only a small ratio of rate constants separates a sequential from an intrasite processive event. To address what a truly sequential process might be, the scenario $k_2=.1k_1$ was modeled, showing a large decrease in activity in comparison to the tritium data and the $k_1 \sim k_2$ scenario.

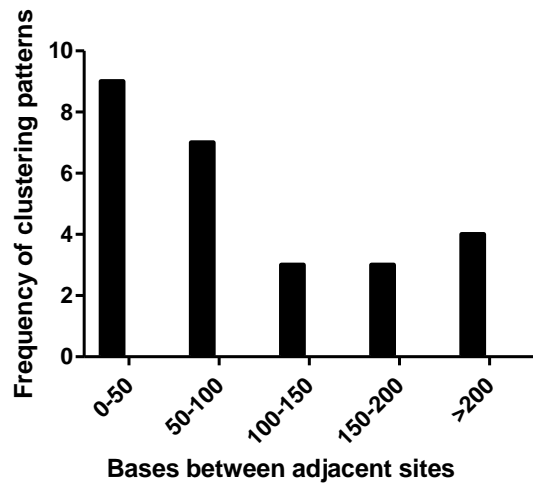
Supplemental Figure 4: (A) Representative gel for the reaction time course for substrate 1B. The DpnI data is derived from the accumulation of the 115mer base pair fragment. The final lane represents the plateau level from the complete methylation of the substrate. (B) Representative gel for substrate 1C.

SUPPLEMENTAL TABLES

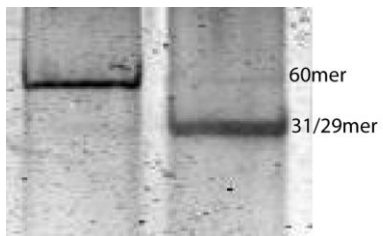
Table 1

number; genome location ^a	distance between sites	sequence
2; 141280-141467	168	TAAACATGTGT GATC GTGCATCACAATTCGAGCTTTATTAACAGATTCGCGGAATGAATAGTTTTCTGGTATACTGCGTGCTTTGCGCTTTGTTGCGGTGCCAAAACCTGCCCGTGCGAAGTGATTTGT TTTTAAATCATATGGTTAGAGATATGAAACATACTGTAGAAGTAAT GATC CCCCGAAGCGG
3; 344331-344419	70	GATATTTTGTGT GATC GGCGACACTTCGCTTAAAAAGCACCAGTAGTGGTTTCGCAGCCATG CGGTGTATAAAAAAT GATC TCATGC
4; 621442-621662	197	TCGTTAT GATC TTATTTGGATATGTTAGCATGTGCAGCCTAAGAATAGGTATTTAAAAATTT GATGGCAAGGCATTGTAATGAATAAACAAATCCTGGCTGCTTAACCTCAGCCTGTTGAAAACGCA CCCGCGCTTTCGCGCAGTATTCCTCGCTCGTTTCATCTCAATTGTGTCTCTGGGTTTGTCTCGGC GTCGCGGTGCCGGTGCA GATC CAGATGAT
5; 765190-765268	55	GTGAAAT GATC CACATAATGGTATTGTTTTATCGGGCACACTGGCGGACTATAAAAA GATC CA AGTGAG GATC ATGAT
6; 1099413-1099492	53	AATAAGTGT GATC TACGTCACTCATAACTGCAACGGATAATTTGTTGTTGCATAAAAATGTTGTG TC GATC TCATTCATGG
7; 1168208-1168301	20	TGTATGT GATC CAGATCACATCTATCATTTAGTTAT GATC GTTAAGTAATTGCTTGCAGCGTC ATTCACTGCATAAGGCCACTATTATGAAA
8; 1365999-1366177	169	TTCAATCA GATC TTTATAAATCAAAAAGATAAAAAATGGCACGAAAATGTTATTAACAGTTCA GCAGGACAATCCTGAACGCAGAAATCAAGAGGACAACATTATGGGTATTTTTTCTCGCTTTGCC GACATCGTGAATGCCAACATCAACGCTCTGTTAGAGAAAAGCGGAA GATC CA
9; 1653141-1653271	23	AACTTCAGCCTTTA GATC TTTCATGTTTCGATTCTTGCATCGCTTGCCTGATGCATGAAATC TACGCAACTGAGCTACTACCATAAAGTATAAA GATC GAAAAAAGCCGGAGT GATC CAAAAAA AGG
10; 1859438-1859479	23	AAAAACACAGATAAT GATC TGCGTTTTACAAC TGATC ACAAA
12; 2069331-2069399	22	AGAATAAAAC GATC AAATATCTATTTTAT GATC GTTTATAT GATC GATAAGCTAATAATAACC TTTGT
13; 2229690-2229814	105	TGATC GCACCGTTCCTTTTCCCGATTATTCTGGCAGTAATGGGCTAAAATTTGCGATGCGTCGC GCATTTTGTATGATGTTTACAGCGTTGCATAATTAATGAGATTC GATC CACATATAAAGC
14; 2599023-2599032	2	CTCGT GATC AG GATC ACA
16; 3490434-3490522	22	GAC GATC ACTTTTTATTCCGT GATC AAAATCACCTCTTAAATGCAATTTAGCAACCGATTGCA ATAAAACATTTAAAC GATC ACAAA
17; 3638634-3638751	96	GTTACCT GATC AGCGTAAACACCTTATCTGGCCTACGGTCTGCGTACGCAATCAAAATCCCCAG CCAATACAACATTTAACACCATCATATTTCCATCATTAGTGT GATC ATCTGGT
18; 3740597-3740743	50	AC GATC TCGACAGATAATTTATAACCAATTGATTTTATGTCTTTTGAATTCATCAATCAGAT TGCCTTGTAAAAAGT GATC GATATATTGAAATCAAGTTTCGCATATTGAAATTTAAGCCAA AAAAG GATC AAAAAAACA
19; 3769939-3770003	39	AACATGCTGT GATC CACATCAGGTGAACGCCGTAAGAAAATATCTTGTGATTC GATC ACAAAAG A
20; 3873223-3873339	98	CCC GATC GATGCATAACCGCGCACTTTGTAGTACCAGCGTGATGACGTTGCGGTTTCCCGTGC GTGTAATGTAGTACAACTTATATTGTTGTACTACAATTT GATC ACAAAAAG
21; 4071646-4071805	133	TGCTTGCTGTTTTT GATC GATATTGTAATTTATCGTCAAAAAATGACAGCCGTCATTTTTTA ACAATTTGGTGAATTTAAATGAACGCATCCCAAAATGTTTAAGGAATGACCATGATTCGTGTT GCTTGTGTAGGTATAACCGTGAT GATC GCAT
22; 4099532-4099618	38	TTTTTGT GATC CAATTTCAAAATAAAAACAAT GATC CGATAAAAAATAAACAGCGTTTCAATTGA TGTGGTTTT GATC ACTTTTTATTG
24; 4347090-4347259	138	AGATTAATCT GATC TACCCATTTGTGGGTA AAAATACACATAATGCGGGTGACATAATAGTTAA TTAACTTTTGTAGCGTTTTGAAATTA AAAACACCGTTCACTGAAGAGATATTAATTTTAGC GATGATGGAGGATAATTTATTT GATC TGGCACAAGTTTTA
25; 4537910-4538119	59	ACCTGTTATACCA GATC AAAAATCACGCAATCCATACAAACAAAACAGATTTGCAATTCGTGTC ACAAAATATGT GATC TTTTTCTAAGAGGAAGATGCCATGTGAAGCCAGACGAACACTTGCAGT GGTCTTCAAAAATAAA GATC TTAGTTTAACTATTTGT
26; 4537959-4538108	65	TTATACCA GATC AAAAATCACGCAATCCATACAAACAAAACAGATTTGCAATTCGTGTCACAAA ATATGT GATC TTTTTCTAAGAGGAAGATGCCATGTGAAGCCAGACGAACACTTGCAGTGGTCT TCAAAAATAAA GATC TTAGTT

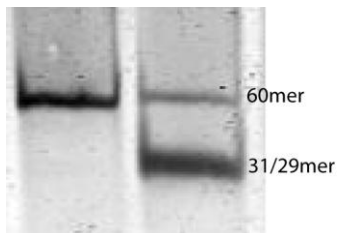
Supplemental Figure 1



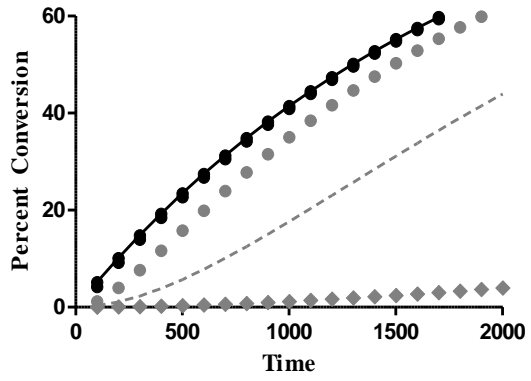
Supplemental Figure 2A:



Supplemental Figure 2B:

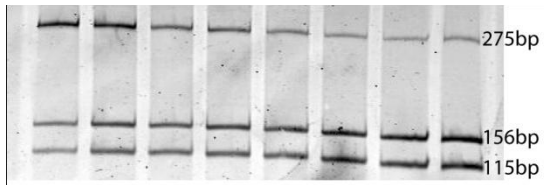


Supplemental Figure 3:

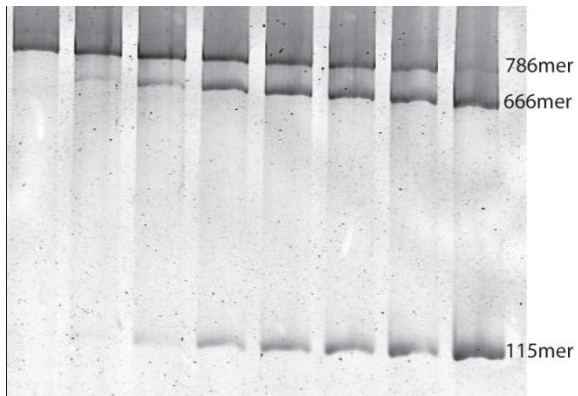


Supplemental Figure 4:

A



B



REFERENCES

1. Casadesus, J., and Low, D. (2006) *Microbiol. Mol. Biol. Rev.* **70**(3): 830-856