

Supplementary Data

Supplementary Table 1. Oligonucleotide sequences

Oligonucleotides used for small RNA hybridization were modified with locked nucleic acid (LNA) bases indicated by a “+” preceding the modified position.

Oligonucleotide Number	Sequence
JP2410	5'-TGGTTCACGTAGTGGGCCATCG-3'
JP3483	5'-GATCTTTGCCGAAAACAATTGGAGG-3'
JP3484	5'-CGACTTGTCATTAGAAAGAAAGAGAT-3'
JP3549	5'-TTAGTTCCGTACCTGATGATGTCG-3'
JP3550	5'-TGAGCTTTTTCTCAACTCTTAAAGG-3'
JP3551	5'-TAAGTTTTATTATTTGGATTAAAGYGGATAATATTATTTAAATYGAAAAAG TTGGAATGGGTTTGGAGAGTTT-3'
JP3552	5'-AATCTCTAAAATTTTTTATTTTACTCATTCTACTTTAAACCTCTATCTTATAA ATACACACAAACCCTAATATATTTTATATTAATAAAA-3'
JP3630	5'-TGTGGGCATTTCAGTTTATAGAGGATTGCCG-3'
JP3631	5'-CTATACCAAAAATGAAGATGAAAGGGAAGG-3'
JP3632	5'-ATATCCGCTCGAGATGGGTGATTGGTCAAAGCTACCAGAGG-3'
JP3633	5'-GCGGGATCCTTAGAGAAAGCTAGGGACAAACATTTGG-3'
JP3634	5'-GCTCTAGAAAGACTCACTGTCTTTCATGTCTCTGC-3'
JP3635	5'-CCGGTGAACAGCTCCTCGCCCTTGCTCACCATTTCTCTCCCCTGTTTTTGCTA CTATTG-3'
JP3636	5'-CAATAGTAGCAAAAACAGGGGAGAGAAATGGTGAGCAAGGGCGAGGAGCT GTTACCCGG-3'
JP3637	5'-GCTCTAGAAGGTCCTGGATTTTGGTTTTAGGAATTAG-3'
JP3687	5'-GAAAACAATAAATTTATTTGCAAACGGGATCCAAATCATCGTACAAAACAT TCAAGATATATGATG-3'
JP3688	5'-CATCATATATCTTGAATGTTTTGTACGATGATTTGGATCCCGTTTTGCAAATAA ATTTATTGTTTTTC-3'
JP5776	5'-GTATAGAGGTTTAAAGTAGAATGAGTAAAATAAAAAAATTTAGAGATTTAY GTYGTTTGGGAAGAATTTT-3'
JP5777	5'-ATACTAAATTTACATTTATATCRATAAATTTAATCRAAAAATATATCTAAAA CCRCAATTTATAATATTAATAAAAAAATAATTAATAA-3'

SDC tandem	
LNA probe	5'-T+GAA+ATCTT+ATCTTT+ATA+ACT+GAC+GTGT+ACT+GTGA+AATT-3'
SDC spreading	
LNA probe #1	5'-A+GG+GTTT+GAT+GTGT+ACCT+ATAA+GT+ATA+GAG+GT-3'
LNA probe #2	5'-C+AC+GCC+GCTT+GGA+AGA+ATTCC+AAAC+ATT+TC-3'
LNA probe #3	5'-+GTT+GTTT+AG+GTTT+ATG+GTGTTT+GG+ATTG-3'

Supplementary Table 2. Sodium bisulfite sequencing of *SDC* – Bisulfite-I

Sodium bisulfite sequencing data for *SDC* with the number of independent clones analyzed shown for each genotype. The % methylation detected for each sequence context is shown, with the number of sites scored listed in parentheses (methylated sites/total sites). The data is analyzed for the entire sequenced region (Total – 10 CG sites, 11 CHG sites, 29 CHH sites), the tandemly repeated region alone (Tandem - 8 CG sites, 11 CHG sites, 17 CHH sites) or the upstream, non-repeated region alone (Spread – 2 CG sites, 0 CHG sites, 12 CHH sites). The *gSDC* transgene restriction site polymorphism (ACATGA to GGATCC) introduces an additional CHH site into the upstream, unique region.

Genotype	# clones	Total %CG	Total %CHG	Total % CHH	Tandem %CG	Tandem %CHG	Tandem %CHH	Spread %CG	Spread %CHH
Col	21	83.8% (176/210)	68.8% (159/231)	10.2% (62/609)	85.7% (144/168)	68.8% (159/231)	10.3% (37/357)	76% (32/42)	9.9% (25/252)
<i>drm1 drm2</i>	20	88% (176/200)	63.6% (140/220)	4% (24/580)	88.8% (142/160)	63.6% (140/220)	1.2% (4/340)	85% (34/40)	8.3% (20/240)
<i>nrpd2a nrpd2b</i>	25	94% (235/250)	78.2% (215/275)	2.3% (17/725)	93.5% (187/200)	78.2% (215/275)	0.2% (1/425)	96% (48/50)	5.3% (16/300)
<i>cmt3</i>	22	72.3% (159/220)	17.8% (43/242)	10.2% (62/609)	14.7% (94/638)	17.8% (43/242)	14.4% (54/374)	63.6% (28/44)	15.2% (40/264)
<i>kyp</i>	18	86.7% (156/180)	42.9% (85/198)	7% (37/522)	88.2% (127/144)	42.9% (85/198)	10.1% (31/306)	80.1% (29/36)	2.8% (6/216)
<i>met1</i>	22	0.03% (6/216)	49.6% (120/242)	12.3% (79/638)	3.4% (6/176)	49.6% (120/242)	20.9% (78/374)	0% (0/40)	0.4% (1/264)
<i>drm1drm2cmt3</i>	22	0% (0/216)	0% (0/242)	0.15% (1/638)	0% (0/176)	0.0% (0/242)	0% (0/374)	0% (0/44)	0.37% (1/264)
F ₁	28	84.3% (236/280)	62.7% (193/308)	14.5% (118/812)	84.3% (189/224)	62.7% (193/308)	17% (81/476)	83.9% (47/56)	11% (37/336)
<i>gSDC Col</i>	30	77.7% (233/300)	72.4% (239/330)	23.4% (211/900)	76.3% (183/240)	72.4% (239/330)	23.9% (122/510)	86.7% (52/60)	22.8% (89/390)
<i>gSDC drm1drm2</i>	30	0% (0/300)	0.6% (2/330)	0% (0/900)	0% (0/240)	0.6% (2/330)	0% (0/510)	0% (0/60)	0% (0/390)

Supplementary Table 3. Sodium bisulfite sequencing of *SDC* – Bisulfite-II

Sodium bisulfite sequencing data for *SDC* with the number of independent clones analyzed shown for each genotype. The % methylation detected for each sequence context is shown, with the number of sites scored listed in parentheses (methylated sites/total sites). The region sequenced contains 1 CG site, 3 CHG sites and 25 CHH sites.

Genotype	# clones	Total %CG	Total %CHG	Total % CHH
Col	29	86.2% (25/29)	55.2% (48/87)	5.5% (40/725)
<i>met1-3</i>	34	0% (0/0)	3.4% (3/102)	0.97% (7/850)

Supplementary Table 4. Characteristics of *SDC* siRNAs

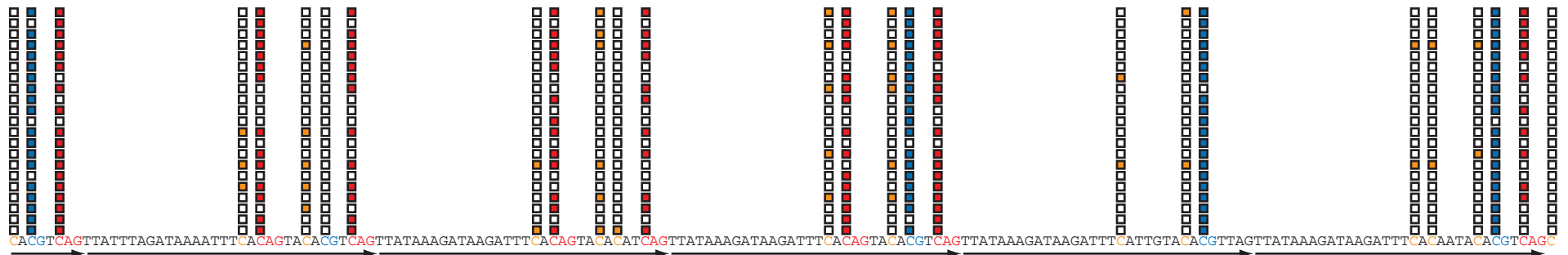
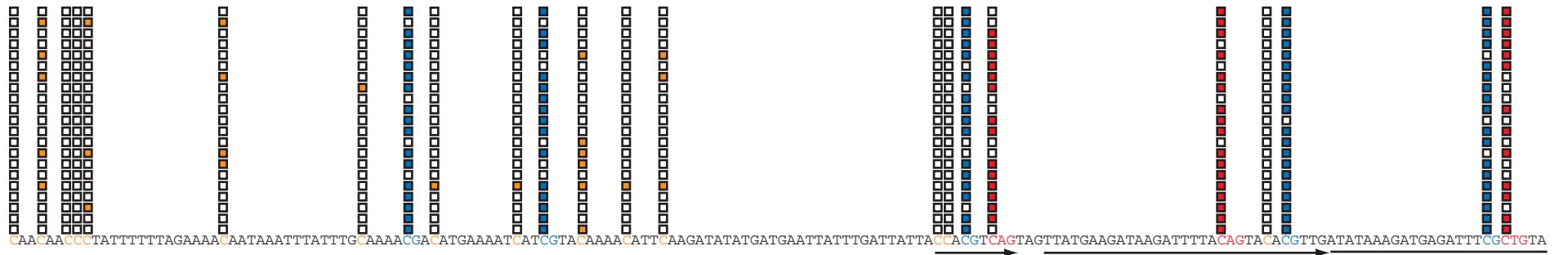
The sequences of siRNAs mapping to the *SDC* region are listed together with information on length and which strand they map to. If the siRNA sequence is unique within the genome it is listed as “U”. Several of the siRNA deriving from the tandem repeats map to one or more of the repeats and are listed as “N”. None of the siRNA mapped to genomic regions other than *SDC*. Whether siRNA mapped to the tandem repeat (T) or spreading (S) region is also listed. Phase was calculated beginning at the first nucleotide of the tandem repeat region (chromosome 2, coordinate 7689171) and assumes that all siRNAs generated in this region are 24nt in length. All siRNA sequences were described in Rajagopalan, et al., (2006).

Sequence	Length (nt)	Strand	Phase	Unique	Spread/Tandem
AAGACATCTCTAAAAAAGTCACCA	24	-	18	U	S
ACCAATGAGAATGCTCTTAGGAAC	24	+	22	U	S
GGTTTTGTAGACTAATTTTCTTTT	24	-	21	U	S
CTTAATTTATCAATTTCCGGTACA	23	-	3	U	S
ACGTTGATATAAAGATGAGATTC	24	+	11	U	T
TAGATAAAAATTTACAGTACACG	23	-	16	U	T
AATTTACAGTACACGTCAGTTAT	24	-	9	U	T
TCACAGTACACGTCAGTTATAAA	23	+	5/13	N	T
CACAGTACACGTCAGTTATAAAGA	24	-	4/12	N	T
AGTACACGTCAGTTATAAAGATAA	24	+	1/9	N	T
GATAAGATTTACAGTACACATCA	24	-	6	U	T
TAAGATTTACAGTACACATCAGT	24	-	4	U	T
GATAAGATTTACAGTACACATCA	24	-	6	U	T
TAAGATTTACAGTACACATCAGT	24	-	4	U	T
GATTTACAGTACACATCAGTTAT	24	-	1	U	T
TTCACAGTACACATCAGTTATA	22	-	22	U	T
TACACATCAGTTATAAAGATAAG	23	-	15	U	T
TCACAGTACACGTCAGTTATAAA	23	+	5/13	N	T
CACAGTACACGTCAGTTATAAAGA	24	-	4/12	N	T
AGTACACGTCAGTTATAAAGATAA	24	+	1/9	N	T
TGTACACGTTAGTTATAAAGATAA	24	+	1	U	T
GATTTACAATACACGTCAGCCCT	24	-	1	U	T

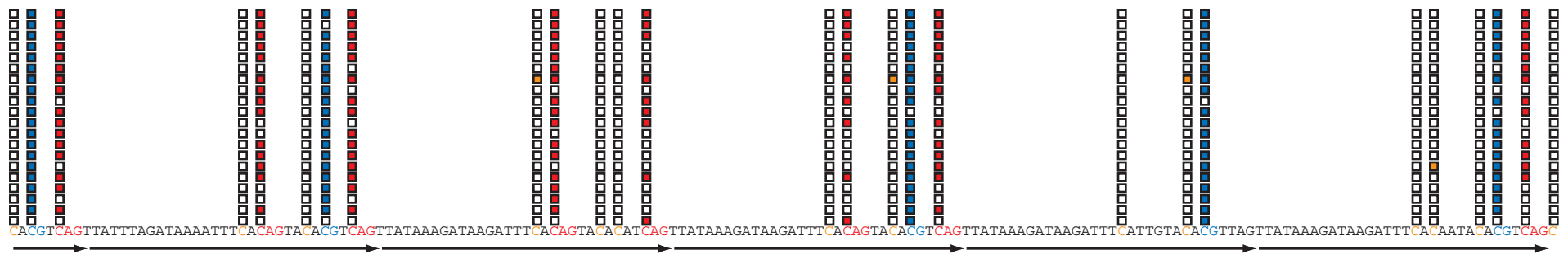
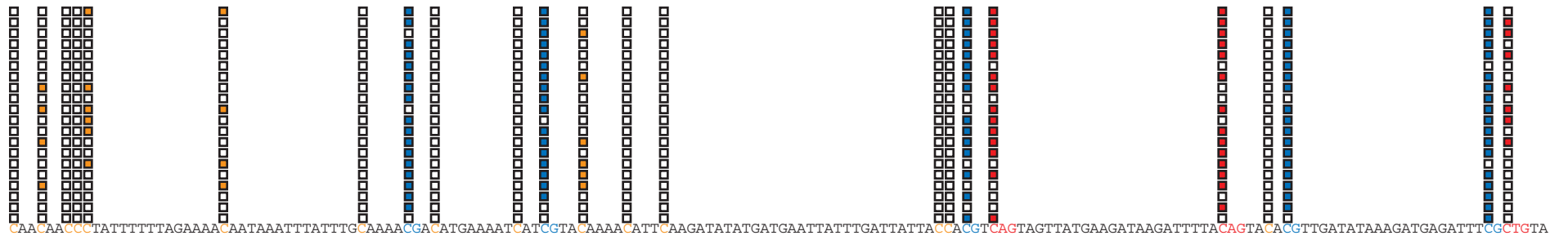
ACAATACACGTCAGCCCTAACACA	24	+	7	U	T
ACATACTAGGGTTTGATGTGTACC	24	+	9	U	S
TTTGATGTGTACCTATAAGTATAG	24	+	20	U	S
AGAGACTCAGCCGCTTGAAGAA	24	+	8	U	S
AATCCAAACACCATTAAAACCTAT	24	-	18	U	S
ACCGATATAAATGTAAATCTAGC	23	+	17	U	S

Supplementary Figure 1. Sodium bisulfite sequencing of *SDC*

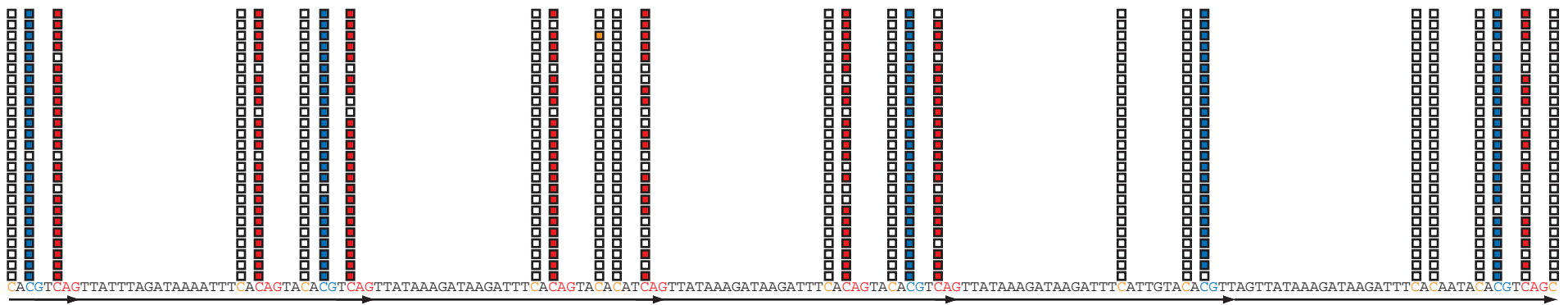
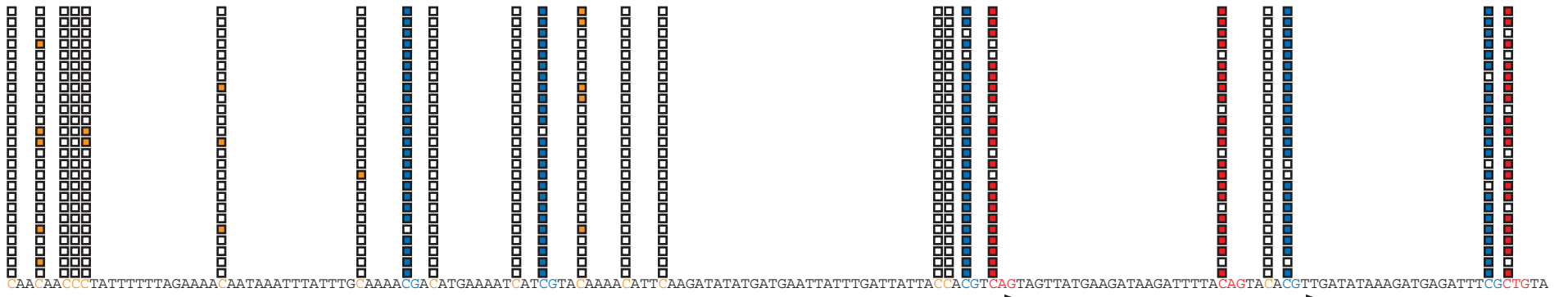
Graphical representation of sodium bisulfite sequencing data at *SDC* in wild type and mutant backgrounds. The primary sequence is listed with cytosines highlighted in blue (CG), red (CHG) and orange (CHH). Tandem repeat regions are highlighted by the arrows underneath the sequence. The consensus sequence for the tandem repeats is AGTTATAAAGATAAGATTTACAGTACACGTC. There is also a short sequence with homology to the repeats immediately prior to the first repeat, which is highlighted by the shorter arrow. Above each cytosine are stacks of square, with each row of squares representing an independently sequenced clone. Shaded boxes represent detected methylation and are color coded as above. The *gSDC* transgene restriction site polymorphism (ACATGA to GGATCC) is highlighted by purple sequence where appropriate. All figures show the Bisulfite-I region, unless indicated by labeling for Bisulfite-II.



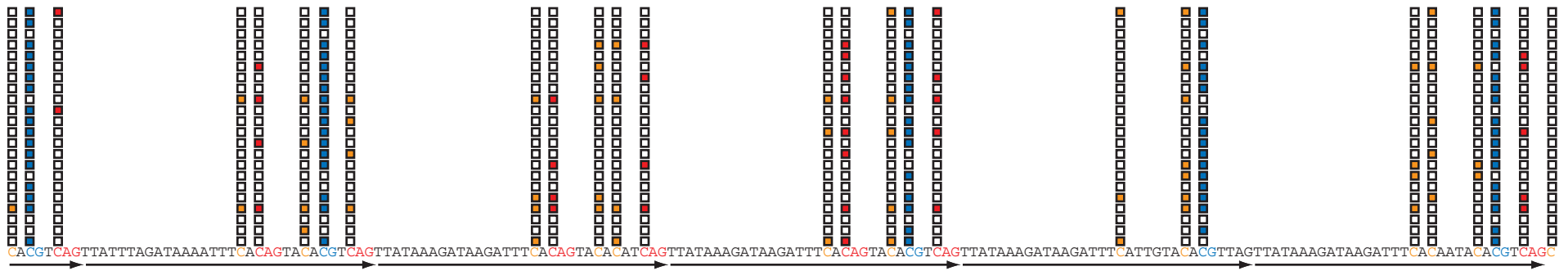
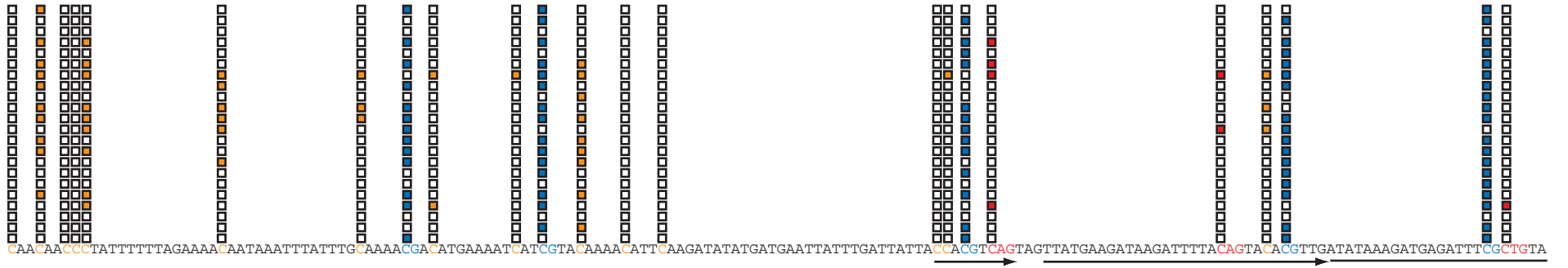
Col



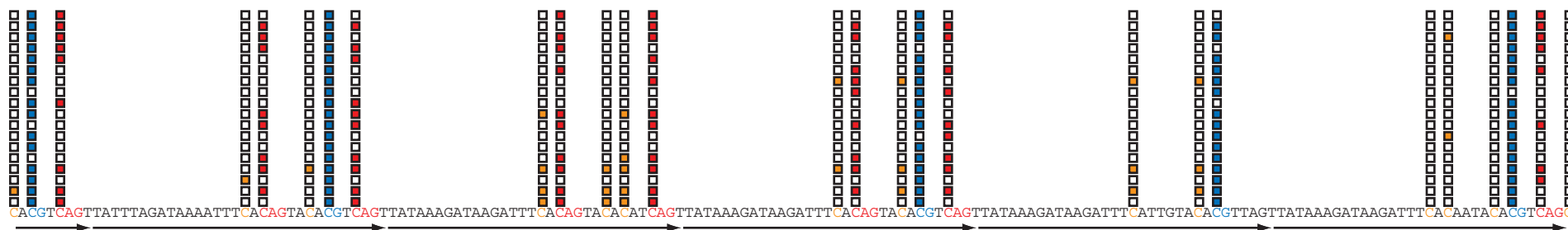
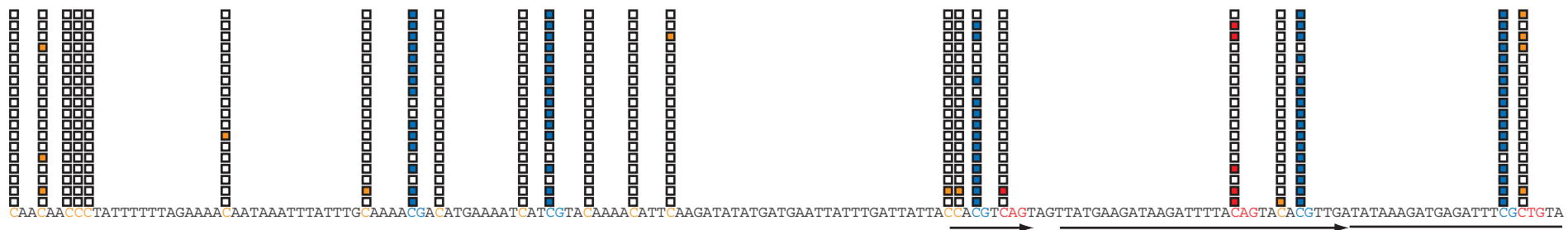
drm1drm2



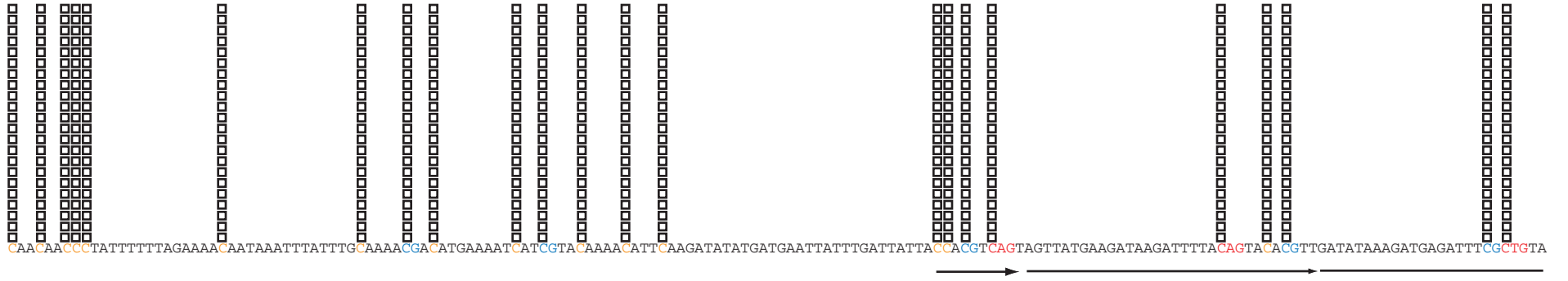
nrpd2a nrpd2b



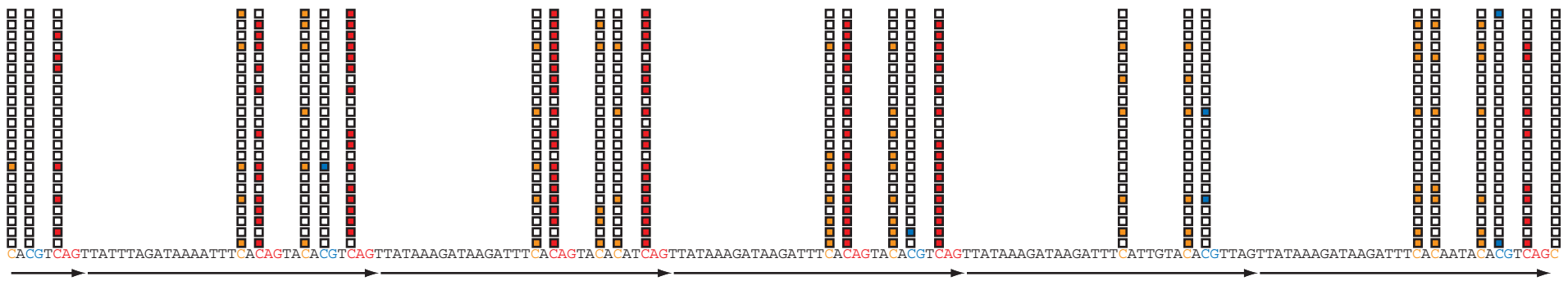
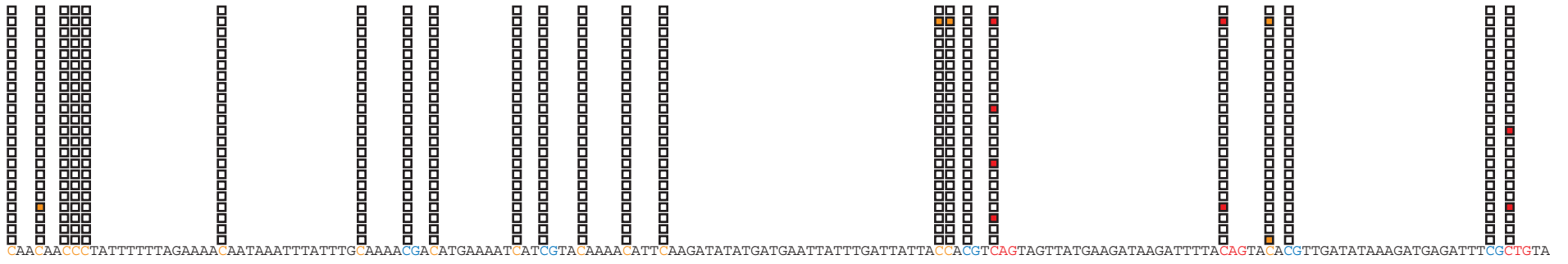
cmt3



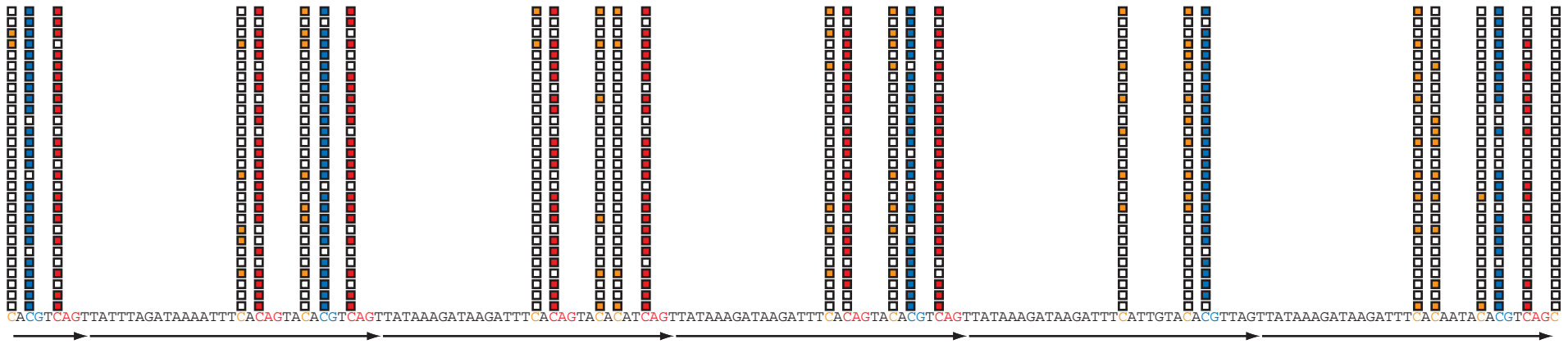
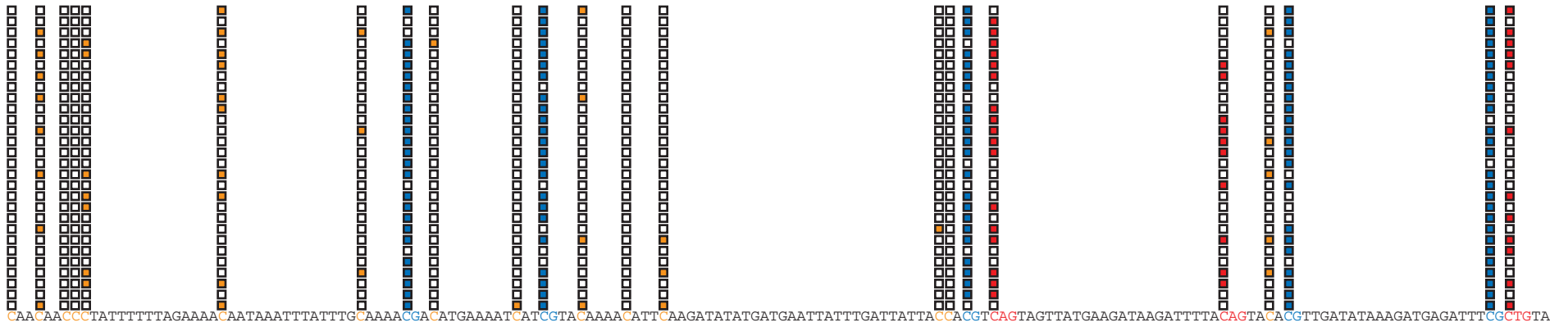
kyp



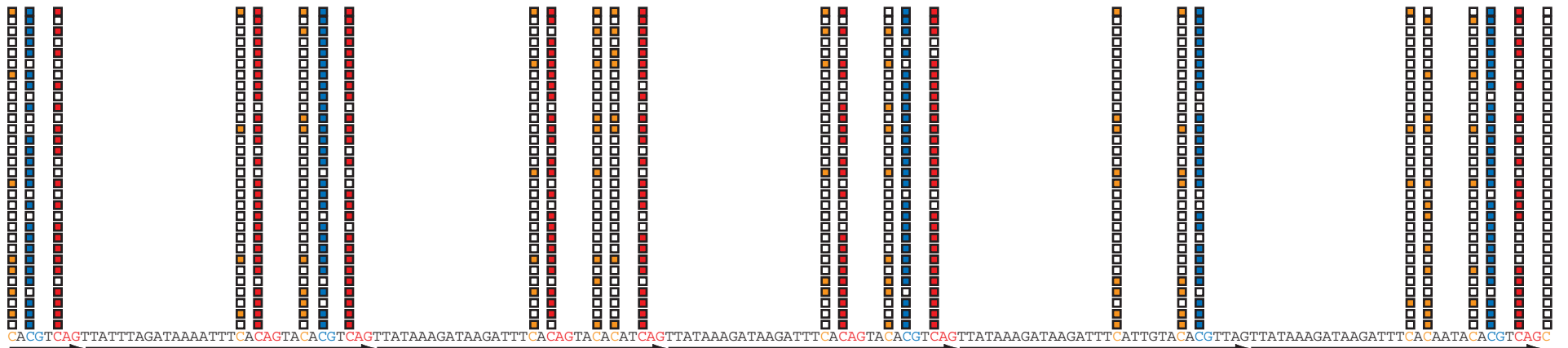
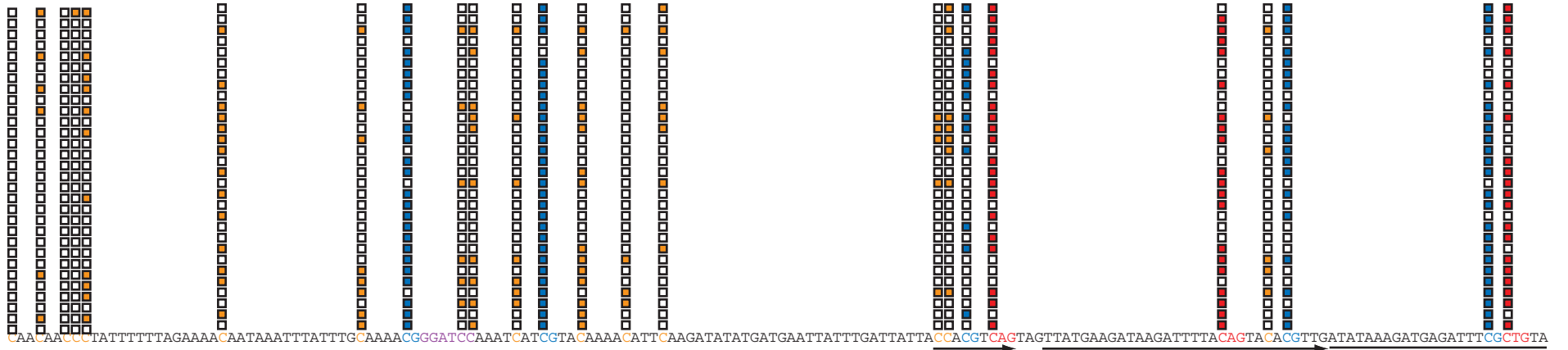
drm1drm2cmt3



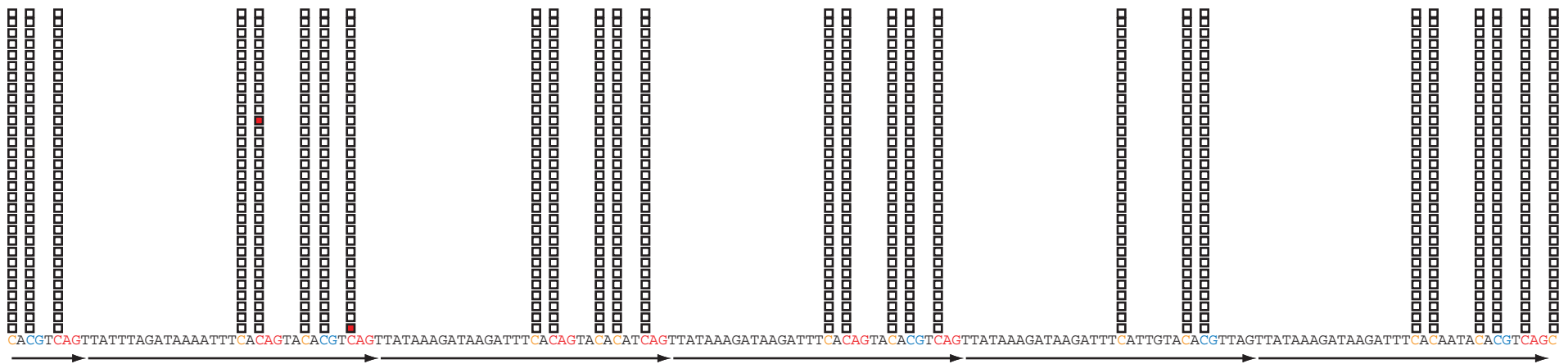
met1



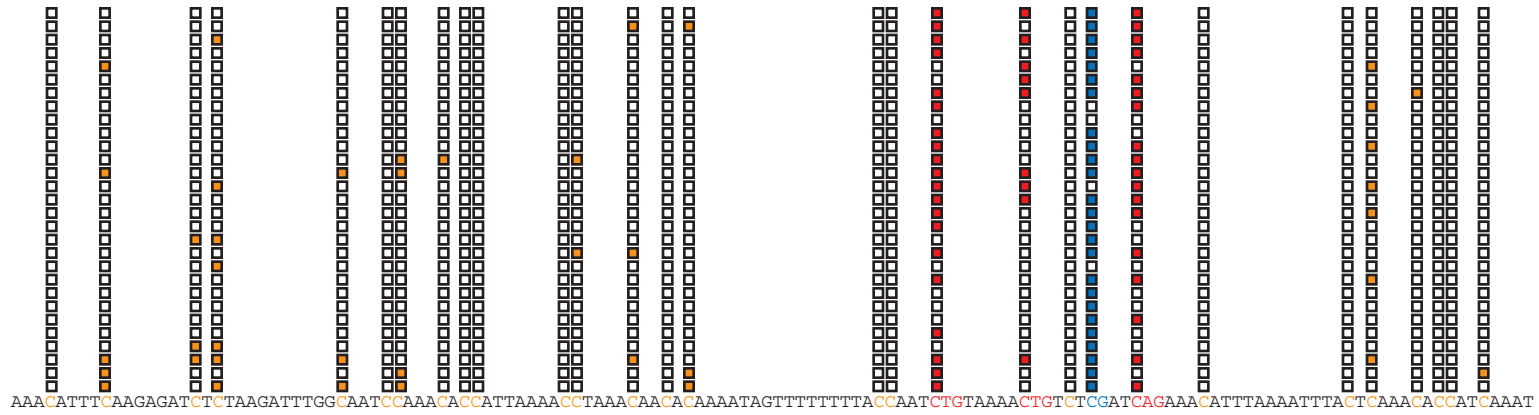
F1 Col x *drm1drm2cmt3*



Col + *gSDC*



drm1drm2 + gSDC



SDC Bisulfite-II Col

