

Table S3. Count breakdown of unique pentamer entries in *methyl*-DNAShape Pentamer Query Table (*mPQT*).

Regular DNA alphabet $\Sigma = \{A, C, G, T\}$			DNA alphabet with two additional letters $\Sigma = \{A, C, G, T, m, g\}$ m: 5mC; g: G base-paired to 5mC			
Strand orientation	5-mers	Count	5-mers containing "mg"	Count	5-mers beginning with "g" or ending with "m"	Count
Forward	NNANN	256	<u>mg</u> NNN	64	NNNNm	256
	NNCNN	256	N <u>mg</u> NN	64	<u>mg</u> NNm	16
			<u>mg</u> <u>mg</u> N	4	N <u>mg</u> Nm	16
			* <u>mg</u> N <u>mg</u>	4	NN <u>mg</u> m	16
					<u>mg</u> <u>mg</u> m	1
					*gNNNm	64
					<u>mg</u> <u>g</u> Nm	4
					gNNNN	256
Reverse	NNTNN	256	NNN <u>mg</u>	64	gNN <u>mg</u>	16
	NNGNN	256	NN <u>mg</u> N	64	gN <u>mg</u> N	16
			N <u>mg</u> <u>mg</u>	4	<u>mg</u> <u>g</u> NN	16
					<u>mg</u> <u>mg</u> <u>g</u>	1
					gNm <u>mg</u> m	4
Total		1024		268		682
Strand-specific total		512		134		341

Total count of pentamers used in the query table = 512 + 134 + 341 = 987

*Only counted in forward strand